GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:50:59 ; Search time 178 Seconds
(without alignments)
5230.639 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggcccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 830525 seqs, 258052604 residues

Total number of hits satisfying chosen parameters: 1661050

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095454_11018/app_query.fasta_1.1991
-DB=SPTREMBL_23 -QFMT=fastan -SUFFIX=rspt -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09806194_@CGN_1_1_256_@runat_25092003_095454_11018 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SPTREMBL_23:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*
15: sp_rvirus:*
16: sp_bacteriap:*
17: sp_archeap:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query % Match	Length	DB	ID	Description
1	2405	72.0	514	11	Q8C5E9	Q8c5e9 mus musculu
2	2399	71.8	514	11	Q8C793	Q8c793 mus musculu
3	2395	71.7	514	11	Q9JL18	Q9jl18 mus musculu
4	2375	71.1	468	4	Q9NZL2	Q9nzl2 homo sapien
5	2293	68.7	439	4	Q9H2V8	Q9h2v8 homo sapien
6	2184	65.4	423	4	Q8N2D4	Q8n2d4 homo sapien
7	1999	59.9	396	4	Q9NZL1	Q9nzl1 homo sapien
8	1246	37.3	255	11	Q9R1P7	Q9r1p7 mus musculu
9	1214.5	36.4	532	4	Q9ULS1	Q9uls1 homo sapien
10	1186	35.5	501	11	Q8C7R1	Q8c7r1 mus musculu
11	1183	35.4	501	11	Q8BQY4	Q8bqy4 mus musculu
12	1172.5	35.1	501	4	Q8IYC8	Q8iyc8 homo sapien
13	1049	31.4	467	11	Q8C4F4	Q8c4f4 mus musculu
14	653	19.6	267	11	Q9CUU5	Q9cuu5 mus musculu
15	461	13.8	213	4	Q9P0D2	Q9p0d2 homo sapien
16	386	11.6	244	5	Q8WQY9	Q8wqy9 aphrocallis
17	367.5	11.0	383	13	Q9DEC3	Q9dec3 xenopus lae
18	361.5	10.8	389	6	Q9GMY4	Q9gmy4 sorex ungui
19	355.5	10.6	384	13	Q91322	Q91322 rana catesb
20	355	10.6	389	13	Q9W643	Q9w643 gallus gall
21	355	10.6	389	13	Q9PWK1	Q9pwk1 gallus gall
22	351.5	10.5	389	6	Q9GMY3	Q9gmy3 rhinolophus
23	351.5	10.5	389	6	Q9GMY5	Q9gmy5 suncus muri
24	335.5	10.0	388	6	Q9GMY2	Q9gmy2 oryctolagus
25	334.5	10.0	391	5	Q9VKP6	Q9vkp6 drosophila
26	326	9.8	399	13	O93458	O93458 podarcis si
27	324.5	9.7	372	5	Q9VLK3	Q9vlk3 drosophila
28	324.5	9.7	383	13	Q9DE45	Q9de45 salvelinus
29	322.5	9.7	390	6	Q8SQ41	Q8sq41 canis famil
30	322.5	9.7	397	13	Q9W6D4	Q9w6d4 hynobius le
31	320	9.6	385	13	Q9DEC4	Q9dec4 rana catesb
32	319.5	9.6	387	13	Q9DDV5	Q9ddv5 salvelinus
33	319	9.6	419	5	Q95VA2	Q95va2 clonorchis
34	318.5	9.5	396	13	O93428	O93428 chionodraco
35	316	9.5	378	13	Q9PUR9	Q9pur9 pseudopleur
36	316	9.5	392	11	Q9D7R7	Q9d7r7 mus musculu
37	313	9.4	383	5	O76856	O76856 dictyosteli
38	312.5	9.4	354	5	Q9GYX7	Q9gyx7 boophilus m
39	305	9.1	384	13	Q9DEC2	Q9dec2 xenopus lae
40	302	9.0	398	13	P87370	P87370 oncorhynchu
41	301.5	9.0	401	11	Q91X66	Q91x66 mus musculu
42	300.5	9.0	386	6	Q9BGU5	Q9bgu5 bos taurus
43	297.5	8.9	387	6	Q9GMY8	Q9gmy8 sorex ungui
44	297	8.9	388	4	Q8N1E3	Q8n1e3 homo sapien
45	295.5	8.8	390	6	Q9GK10	Q9gk10 camelus dro

ALIGNMENTS

RESULT 1
Q8C5E9
ID Q8C5E9 PRELIMINARY; PRT; 514 AA.
AC Q8C5E9;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Beta-site APP-cleaving enzyme 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Testis;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK078770; BAC37384.1; -.
SQ SEQUENCE 514 AA; 55811 MW; CBB9237BB68A0B2E CRC64;

Alignment Scores:

Pred. No.:	2.16e-177	Length:	514
Score:	2405.00	Matches:	460
Percent Similarity:	92.66%	Conservative:	20
Best Local Similarity:	88.80%	Mismatches:	34
Query Match:	72.03%	Indels:	4
DB:	11	Gaps:	1

US-09-806-194-1 (1-1804) x Q8C5E9 (1-514)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||   |||||||||||||||||:::|||||||||||||||||||||||||   |||
Db          1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuProValLeuAlaGlnTrpLeuLeuSerAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
                 |||   ||||||||||||||||||||||||||||:::||||||    .||||||||
Db         21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaArgAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
                  :::      ||||||   ||||||||          ||||||||||||
Db         41 HisArgAlaSerAlaValProGlyLeuGlyThrProGlyLeuProArgAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||      |||:::            ||||||||||||||||
Db         61 AlaLeuAlaLeuGluProValArgAlaThr----------AlaAsnPheLeuAlaMet 76

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 96

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
                 ||||||||||||||||||||||||||||:::|||   ||||||||||   |||
Db        117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||:::|||||||||||||||||||||||||||
Db        157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||   ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||   ||||||:::::|||||||||||||||||||||||||||||||||||||||
Db        217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 276

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||||
Db        277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||   ||||||||||||||||||||||||||||||||||||||:::
Db        317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrProLysIleSerIle 356

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||
Db        357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 376

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||   ||||||||||||||||||||||||||||||||||
Db        377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||:::||||||
Db        397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              |||||||||:::||||||||||||||   ||||||||||||||||   |||
Db        417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||:::||||||||||||||||||||||
Db        437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              |||:::|||:::||||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GlnAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 476

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||:::||||||||||||||||      |||:::      |||
Db        477 AlaIleLeuLeuValLeuLeuIleLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        497 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 514
```

RESULT 2
Q8C793
ID Q8C793 PRELIMINARY; PRT; 514 AA.
AC Q8C793;
DT 01-MAR-2003 (TrEMBLrel. 23, Created)
DT 01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Beta-site APP-cleaving enzyme 2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Heart;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RA the RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK052309; BAC34931.1; -.
SQ SEQUENCE 514 AA; 55871 MW; 8BF45E07B0990225 CRC64;

Pred. No.:	6.29e-177	Length:	514
Score:	2399.00	Matches:	459
Percent Similarity:	92.47%	Conservative:	20
Best Local Similarity:	88.61%	Mismatches:	35
Query Match:	71.85%	Indels:	4
DB:	11	Gaps:	1

US-09-806-194-1 (1-1804) x Q8C793 (1-514)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC  60
Db          1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuProValLeuAlaGlnTrpLeuLeuSerAla  20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC  120
Db         21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaArgAlaThrAsn  40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG  180
Db         41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu  60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG  240
Db         61 AlaLeuAlaLeuGluProValArgAlaThr----------AlaAsnPheLeuAlaMet  76

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC  300
Db         77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr  96

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA  360
Db         97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC  420
Db        117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA  480
Db        137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT  540
Db        157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGCCCTAGCT  600
Db        177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA  660
Db        197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT  720
Db        217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA  780
Db        237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA  840
Db        257 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 276

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC  900
Db        277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG  960
Db        297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIle 356

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgThrThrIleLeuProGlnLeuTyr 376

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        457 GlnAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 476

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        477 AlaIleLeuLeuValLeuIleLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        497 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 514
```

RESULT 3
Q9JL18
ID Q9JL18 PRELIMINARY; PRT; 514 AA.
AC Q9JL18;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aspartyl protease 1.
GN BACE2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA Choi D.K., Sugano S., Sakaki Y.;
RT "Molecular characterization of the mouse Asp1 gene, a homolog of the
RT human ASP1 (Down Syndrome Region aspartyl protease).";
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AF216310; AAF36599.1; -.
DR HSSP; P00797; 2REN.
DR MEROPS; A01.041; -.
DR MGD; MGI:1860440; Bace2.
DR InterPro; IPR001461; Aspprotease A1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Aspartyl protease; Hydrolase; Protease.
SQ SEQUENCE 514 AA; 55799 MW; A70725F2C1DF5B47 CRC64;

Alignment Scores:

Pred. No.:	1.28e-176	Length:	514
Score:	2395.00	Matches:	459
Percent Similarity:	92.47%	Conservative:	20
Best Local Similarity:	88.61%	Mismatches:	35
Query Match:	71.73%	Indels:	4
DB:	11	Gaps:	1

US-09-806-194-1 (1-1804) x Q9JL18 (1-514)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuLeuValLeuAlaGlnTrpLeuLeuSerAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaGlyAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProValArgAlaThr----------AlaAsnPheLeuAlaMet 76

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 96

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db         97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db        177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db        197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        257 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 276

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIle 356

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 376

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        457 GlnAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 476

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        477 AlaIleLeuLeuValLeuIleLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        497 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 514
```

RESULT 4
Q9NZL2
ID Q9NZL2 PRELIMINARY; PRT; 468 AA.
AC Q9NZL2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Aspartyl protease.
GN BACE2.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20422477; PubMed=10965118;
RA Solans A., Estivill X., de La Luna S.;
RT "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT Alzheimer's amyloid precursor protein beta-secretase.";
RL Cytogenet. Cell Genet. 89:177-184(2000).
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AF188276; AAF35835.1; -.
DR HSSP; P00797; 2REN.
DR InterPro; IPR001461; Aspprotease A1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Aspartyl protease; Hydrolase; Protease.
SQ SEQUENCE 468 AA; 50324 MW; 717E0920126A0142 CRC64;

Alignment Scores:

Pred. No.:	4.39e-175	Length:	468
Score:	2375.00	Matches:	468
Percent Similarity:	90.35%	Conservative:	0
Best Local Similarity:	90.35%	Mismatches:	0
Query Match:	71.13%	Indels:	50
DB:	4	Gaps:	1

US-09-806-194-1 (1-1804) x Q9NZL2 (1-468)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
```

```
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGC: 600
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        321 GluAlaValAlaArgAlaSerLeu---------------------------------- 328

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        328 ---------------------------------------------------------- 328

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140

Db        329 ----------------------------------------------------LeuTyr 330

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        331 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 350

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        351 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 370

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        371 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 390

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        391 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 410

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        411 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 430

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        431 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 450

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        451 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 468
```

RESULT 5
Q9H2V8
ID Q9H2V8 PRELIMINARY; PRT; 439 AA.
AC Q9H2V8;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE CDA13.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Pheochromocytoma;
RA Li Y., Huang Q., Peng, y, Song H., Yu Y., Xu S., Ren S., Chen Z.,
RA Han Z.;
RL Submitted (DEC-1999) to the EMBL/GenBank/DDBJ databases.
CC -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR EMBL; AF212252; AAG41783.1; -.
DR HSSP; P00797; 2REN.
DR InterPro; IPR001461; Aspproteasea1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 1.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Aspartyl protease; Hydrolase; Protease.
SQ SEQUENCE 439 AA; 48275 MW; 02EC0E0E50F11602 CRC64;

Alignment Scores:

Pred. No.:	9.45e-169	Length:	439
Score:	2293.00	Matches:	439
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	68.67%	Indels:	0
DB:	4	Gaps:	0

US-09-806-194-1 (1-1804) x Q9H2V8 (1-439)

```
Qy        238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG 297
Db          1 MetValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGly 20
```

```
Qy        298 ACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA 357
Db         21 ThrProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAla 40

Qy        358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC 417
Db         41 GlyThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArg 60

Qy        418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGCGGCTTCGTTGGG 477
Db         61 SerLysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGly 80

Qy        478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT 537
Db         81 GluAspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThr 100

Qy        538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTA 597
Db        101 IlePheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeu 120

Qy        598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG 657
Db        121 AlaTyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuVal 140

Qy        658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT 717
Db        141 ThrGlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProVal 160

Qy        718 GCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTAT 777
Db        161 AlaGlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyr 180

Qy        778 AAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTG 837
Db        181 LysGlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeu 200

Qy        838 AAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAG 897
Db        201 LysLeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLys 220

Qy        898 GCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTG 957
Db        221 AlaIleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaVal 240

Qy        958 GTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGG 1017
Db        241 ValGluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGly 260

Qy       1018 TCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCC 1077
Db        261 SerGlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSer 280

Qy       1078 ATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTT 1137
Db        281 IleTyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeu 300

Qy       1138 TACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCC 1197
Db        301 TyrIleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSer 320

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db        321 ProSerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePhe 340

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
Db        341 AspArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAla 360

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
Db        361 AlaValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValPro 380

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
Db        381 AlaGlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCys 400

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGC 1497
Db        401 GlyAlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArg 420

Qy       1498 CCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        421 ProArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 439
```

RESULT 6
Q8N2D4
ID Q8N2D4 PRELIMINARY; PRT; 423 AA.
AC Q8N2D4;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE Hypothetical protein OVARC1000363.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Ovarian carcinoma;
RA Ota T., Nishikawa T., Suzuki Y., Kawai-Hio Y., Hayashi K., Ishii S.,
RA Saito K., Yamamoto J., Wakamatsu A., Nagai T., Nakamura Y.,
RA Nagahari K., Sugano S., Isogai T.;
RT "HRI human cDNA sequencing project.";
RL Submitted (MAR-2002) to the EMBL/GenBank/DDBJ databases.
DR EMBL; AK075539; BAC11682.1; -.
DR InterPro; IPR001461; Aspprotease A1.
DR InterPro; IPR001969; Aspprotease_site.
DR Pfam; PF00026; asp; 2.
DR PRINTS; PR00792; PEPSIN.
DR PROSITE; PS00141; ASP_PROTEASE; 2.
KW Hypothetical protein.
SQ SEQUENCE 423 AA; 46457 MW; 4D4839F2ED9C2CE1 CRC64;

Alignment Scores:

Pred. No.:	2.52e-160	Length:	423
Score:	2184.00	Matches:	420
Percent Similarity:	99.53%	Conservative:	1
Best Local Similarity:	99.29%	Mismatches:	2
Query Match:	65.41%	Indels:	0
DB:	4	Gaps:	0

US-09-806-194-1 (1-1804) x Q8N2D4 (1-423)

```
Qy        286 ATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC 345
Db          1 MetLeuIleGlyThrProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsn 20

Qy        346 TTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCT 405
Db         21 PheAlaValAlaGlyThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSer 40

Qy        406 AGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACG 465
Db         41 SerThrTyrArgSerLysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThr 60

Qy        466 GGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTC 525
Db         61 GlyPheValGlyGluAspLeuValThrIleProLysGlyPheAsnThrSerPheLeuVal 80

Qy        526 AACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGA 585
Db         81 AsnIleAlaThrIlePheGluSerGlyAsnPhePheLeuProGlyIleGlnTrpAsnGly 100

Qy        586 ATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTC 645
Db        101 IleLeuGlyLeuAlaTyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePhe 120
```

```
Qy       646 GACTCCCTGGTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCC 705
Db       121 AspSerLeuValThrGlnAlaAsnIleProAsnValPheSerMetGlnMetArgGlyAla 140

Qy       706 GGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAA 765
Db       141 GlyLeuProValAlaGlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGlu 160

Qy       766 CCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAG 825
Db       161 ProSerLeuTyrLysGlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGln 180

Qy       826 ATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTAT 885
Db       181 IleGluIleLeuLysLeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyr 200

Qy       886 AACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTG 945
Db       201 AsnAlaAspLysAlaIleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysVal 220

Qy       946 TTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGT 1005
Db       221 PheAspAlaValValGluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGly 240

Qy      1006 TTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC 1065
Db       241 PheTrpThrGlySerGlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPhe 260

Qy      1066 CCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATC 1125
Db       261 ProLysIleSerIleTyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIle 280

Qy      1126 CTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGA 1185
Db       281 LeuProGlnLeuTyrIleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArg 300

Qy      1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
Db       301 PheGlyIleSerProSerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPhe 320

Qy      1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
Db       321 TyrValIlePheAspArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGlu 340

Qy      1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
Db       341 IleAlaGlyAlaAlaValSerGluIleSerGlyProPheSerThrGluAspValAlaSer 360

Qy      1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC 1425
Db       361 AsnCysValProAlaGlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeu 380

Qy      1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGG 1485
Db       381 MetSerValCysGlyAlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArg 400

Qy      1486 TGTCAGCGTCGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACAT 1545
Db       401 CysGlnArgArgProArgAspProGluValValAsnAspGluSerSerLeuValArgHis 420

Qy      1546 CGCTGGAAA 1554
              \\\\\\\\\
Db       421 ArgTrpLys 423
```

```
RESULT 7
Q9NZL1
ID   Q9NZL1      PRELIMINARY;      PRT;   396 AA.
AC   Q9NZL1;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartyl protease.
GN   BACE2.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20422477; PubMed=10965118;
RA   Solans A., Estivill X., de La Luna S.;
RT   "A new aspartyl protease on 21q22.3, BACE2, is highly similar to
RT   Alzheimer's amyloid precursor protein beta-secretase.";
RL   Cytogenet. Cell Genet. 89:177-184(2000).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR   EMBL; AF188277; AAF35836.1; -.
DR   HSSP; P00797; 2REN.
DR   InterPro; IPR001461; Aspprotease A1.
DR   InterPro; IPR001969; Aspprotease_site.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Aspartyl protease; Hydrolase; Protease.
SQ   SEQUENCE   396 AA;  43013 MW;  5023A7AF391CEAC9 CRC64;
```

```
Alignment Scores:
Pred. No.:                5.02e-146       Length:          396
Score:                    1999.00         Matches:         394
Percent Similarity:       87.42%          Conservative:    2
Best Local Similarity:    86.98%          Mismatches:      0
Query Match:              59.87%          Indels:          57
DB:                       4               Gaps:            1

US-09-806-194-1 (1-1804) x Q9NZL1 (1-396)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
```

```
Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAACTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGln------ 378

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200

Db        378 ------------------------------------------------------------ 378

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260

Db        378 ------------------------------------------------------------ 378

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
                                                  ::: |||:::|||||||||
Db        379 ------------------------------------------Lys-LeuArgValLeuGl 384

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATG 1357
              |||||||||||||||||||||||||||||||||||||
Db        384 nCysLeuLysPheProGlyLeuSerGlnGlnArgMet 396

RESULT 8
Q9R1P7
ID   Q9R1P7      PRELIMINARY;      PRT;   255 AA.
AC   Q9R1P7;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Aspartyl protease (Fragment).
GN   BACE2.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Accarino M., Fumagalli P., Taramelli R., Ottolenghi S.;
RT   "Cloning of a gene from chromosome 21 Down Region encoding a potential
RT   transmembrane protease.";
RL   Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF051150; AAD45964.1; -.
DR   MEROPS; A01.041; -.
DR   MGD; MGI:1860440; Bace2.
DR   InterPro; IPR001969; Aspprotease_site.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Protease.
FT   NON_TER       1       1
SQ   SEQUENCE   255 AA;  28685 MW;  53DE317815996D63 CRC64;

Alignment Scores:
Pred. No.:              7.61e-88        Length:         255
Score:                  1246.00         Matches:        232
Percent Similarity:     95.69%          Conservative:   12
Best Local Similarity:  90.98%          Mismatches:     11
Query Match:            37.32%          Indels:         0
DB:                     11              Gaps:           0

US-09-806-194-1 (1-1804) x Q9R1P7 (1-255)

Qy        790 TGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATT 849
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 TrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLysLeuGluIle 20

Qy        850 GGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGAC 909
              |||||||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db         21 GlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAlaIleValAsp 40

Qy        910 AGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTG 969
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 SerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValValGluAlaVal 60

Qy        970 GCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCG 1029
              ||||||   ||||||||||||||||||||||||||||||||||||:::|||||||||||
Db         61 AlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAlaGlnLeuAla 80

Qy       1030 TGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGA 1089
              ||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db         81 CysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIleTyrLeuArg 100

Qy       1090 GATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCC 1149
              |||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         101 AspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyrIleGlnPro 120

Qy       1150 ATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAAT 1209
              ||||||||||||||   ||||||||||||||||||||||||||||   |||||||||
Db         121 MetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSerSerThrAsn 140

Qy       1210 GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAG 1269
              ||||||||||||||||||||||||||||||||||||||||||||||::|||||||||||
Db         141 AlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAspArgAlaGln 160

Qy       1270 AAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAA 1329
              :::|||||||||||||||   ||||||||||||||||   |||      |||||||||
Db         161 ArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThrValSerGlu 180

Qy       1330 ATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTG 1389
              ||||||||||||||||||||||||||:::|||||||||||||||||||||||||:::|||
Db         181 IleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAlaGlnAlaLeu 200

Qy       1390 AGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTC 1449
              :::||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         201 AsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGlyAlaIleLeu 220

Qy       1450 CTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCCCGTGACCCT 1509
              ||||||||||||:::|||||||||:::|||      |||:::      ||||||||||||
Db         221 LeuValLeuIleLeuLeuLeuLeuValProLeuHisCysArgHisAlaProArgAspPro 240

Qy       1510 GAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              |||||||||||||||||||||||||||||||||||||||||||||
Db         241 GluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 255

RESULT 9
Q9ULS1
ID   Q9ULS1      PRELIMINARY;      PRT;   532 AA.
AC   Q9ULS1;
```

```
DT   01 MAY-2000 (TrEMBLrel. 13, Created)
DT   01-OCT-2001 (TrEMBLrel. 18, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Hypothetical protein KIAA1149 (Fragment).
GN   KIAA1149.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
RX   MEDLINE=20039618; PubMed=10574461;
RA   Hirosawa M., Nagase T., Ishikawa K., Kikuno R., Nomura N., Ohara O.;
RT   "Characterization of cDNA clones selected by the GeneMark analysis
RT   from size-fractionated cDNA libraries from human brain.";
RL   DNA Res. 6:329-336(1999).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
DR   EMBL; AB032975; BAA86463.2; -.
DR   HSSP; P56272; 1AM5.
DR   InterPro; IPR001461; Aspproteasel.
DR   InterPro; IPR001969; Aspprotease_site.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Hypothetical protein; Aspartyl protease; Hydrolase; Protease.
FT   NON_TER       1      1
SQ   SEQUENCE   532 AA;  58720 MW;  98B135D0D5FBD2E8 CRC64;

Alignment Scores:
Pred. No.:              2.64e-85      Length:          532
Score:                  1214.50       Matches:         246
Percent Similarity:     61.81%        Conservative:    81
Best Local Similarity:  46.50%        Mismatches:      179
Query Match:            36.37%        Indels:          24
DB:                     4             Gaps:            8

US-09-806-194-1 (1-1804) x Q9ULS1 (1-532)

Qy          2 TGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTGCTGCGCGCCG 61
              |||         ||||||                |||  |||  ||||         |||
Db          9 TrpGlyGlnAlaProGlyThrAspValGlyGlnCysGluProArgGly---------Pro 25

Qy         62 CCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCC---------GGGTGCCCGCGG 112
                  ||||||   |||       ||||||              |||||   |||
Db         26 GluGlyArgGlyProProTrpProLysProCysProGlySerCysCysGlyTrpAlaArg 45

Qy        113 CCACGAACCGCGTA-GTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCC 171
                       ||:|||:::      |||  |||
Db         46 GluCysCysLeuProThrAlaProSerThrAlaSerGlyCysProCysAlaAlaAlaTrp 65

Qy        172 GACGGCTTGGCGCTCGCCCTG----------------------GAGCCTGCCCTGGCG 207
                |||      |||  |||                      |||||
Db         66 GlyGlyAlaProLeuGlyLeuArgLeuProArgGluThrAspGluGluPro------Glu 83

Qy        208 TCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGC 267
                |||           :::|||:::   |||||||||||||||:::|||   |||||
Db         84 GluProGlyArgArgGlySerPheValGluMetValAspAsnLeuArgGlyLysSerGly 103

Qy        268 CGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTT 327
              :::||||||||||:::||||||   :::|||:::|||||||||   |||   ||||||||
Db        104 GlnGlyTyrTyrValGluMetThrValGlySerProProGlnThrLeuAsnIleLeuVal 123

Qy        328 GACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTAC 387
              ||||||||||||||||||||||||||||    ||||||  ::::::      |||
Db        124 AspThrGlySerSerAsnPheAlaValGlyAlaAlaProHisProPheLeuHisArgTyr 143

Qy        388 TTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTAC 447
              :::      :::   ||||||||||||||||              |||   |||   |||
Db        144 TyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLysGlyValTyrValProTyr 163

Qy        448 ACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTC 507
              |||||||||   |||   |||   :::|||   ||||||||||:::||||||   |||

Db        164 ThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuValSerIleProHisGlyPro 183

Qy        508 AATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCT 567
              |||   :::         |||||||||   |||   |||||||:::   |||||||:::
Db        184 AsnValThrValArgAlaAsnIleAlaAlaIleThrGluSerAspLysPhePheIleAsn 203

Qy        568 GGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGT 627
              |||       |||   ||||||||||||||||||||||||   :::|||:::|||
Db        204 GlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGluIleAlaArgProAspAsp 223

Qy        628 TCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATCCCCAACGTTTTCTCC 687
              |||||||||   ||||||||||||||||||   |||   ::::::||||||:::||||||
Db        224 SerLeuGluProPhePheAspSerLeuValLysGlnThrHisValProAsnLeuPheSer 243

Qy        688 ATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCT---------GGGACCAACGGA 738
              :::|||:::|||||||||||||   |||:::      |||          :::   |||
Db        244 LeuGlnLeuCysGlyAlaGlyProProLeuAsnGlnSerGluValLeuAlaSerValGly 263

Qy        739 GGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACC 798
              ||||||:::::::::|||||||||:::   ||||||||   |||   :::|||||||||
Db        264 GlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThrGlySerLeuTrpTyrThr 283

Qy        799 CCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAA 858
              |||||||::   |||||||||||||:::::   |||:::::::::|||||   ||||||
Db        284 ProIleArgArgGluTrpTyrTyrGluValIleIleValArgValGluIleAsnGlyGln 303

Qy        859 AGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACC 918
                 |||   :::||||||:::|||||||||   ||||||:::||||||||||||||||||
Db        304 AspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSerIleValAspSerGlyThr 323

Qy        919 ACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCA 978
              |||   |||||||||:|||:::|||||||||:::|||   |||:::::::::   |||
Db        324 ThrAsnLeuArgLeuProLysLysValPheGluAlaAlaValLysSerIleLysAlaAla 343

Qy        979 TCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACG 1038
              |||         :::|||   ||||||||||||   |||   ||||||   ||||||
Db        344 SerSerThrGluLysPheProAspGlyPheTrpLeuGlyGluGlnLeuValCysTrpGln 363

Qy       1039 AATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAAC 1098
                    ||||||||||:::   ||||||   ||||||:::||||||      |||
Db        364 AlaGlyThrThrProTrpAsnIlePheProValIleSerLeuTyrLeuMetGlyGluVal 383

Qy       1099 TCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGG 1158
              :::::::::|||||||||||||||||||||||||||||   |||:::::::|:::
Db        384 ThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyrLeuArgProValGluAsp 403

Qy       1159 GCCGGCCTGAATIAT---GAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTG 1215
                    :::      :::|||||||:::|||   ||||||   |||:::
Db        404 ValAlaThrSerGlnAspAspCysTyrLysPheAlaIleSerGlnSerSerThrGlyThr 423

Qy       1216 GTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGG 1275
              |||:::|||||   :::|||||||||||||||||:::|||||||||||||:::||||||
Db        424 ValMetGlyAlaValIleMetGluGlyPheTyrValValPheAspArgAlaArgLysArg 443

Qy       1276 GTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCC 1335
              :::||||||||||   |||   |||                        :::   :::
Db        444 IleGlyPheAlaValSerAlaCysHisValHisAspGluPheArgThrAlaAlaValGlu 463

Qy       1336 GGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAG 1395
              |||||||||   |||   |||:::                              |||
Db        464 GlyProPheValThrLeuAspMetGluAspCysGlyTyrAsnIleProGlnThrAspGlu 483

Qy       1396 CCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTC 1455
                    |||      ::::::|||   :::   ::::::|||   |||:::   ::::::
Db        484 SerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys---AlaLeuPheMetLeu 502

Qy       1456 TTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT------CGCCCCCGTGACCCT 1509
                 :::   |||::::::      :::||||||   |||      |||   :::
Db        503 ProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCysLeuArgGlnGlnHisAsp 522

Qy       1510 GAGGTCGTCAATGATGAGTCCTCTCTG 1536
                 :::      :::|||   |||   |||
Db        523 AspPheAlaAspAspIleSerLeuLeu 531
```

```
RESULT 10
Q8C7R1
ID   Q8C7R1      PRELIMINARY;      PRT;   501 AA.
AC   Q8C7R1;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Beta-site APP cleaving enzyme.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Spinal cord;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK049626; BAC33844.1; -.
SQ   SEQUENCE   501 AA;  55761 MW;  B410DA8B64647663 CRC64;

Alignment Scores:
Pred. No.:              4.14e-83      Length:          501
Score:                  1186.00       Matches:         238
Percent Similarity:     62.14%        Conservative:    82
Best Local Similarity:  46.21%        Mismatches:      169
Query Match:            35.52%        Indels:          26
DB:                     11            Gaps:            7

US-09-806-194-1 (1-1804) x Q8C7R1 (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla-------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrIleGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        307 LysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGlu 326

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePheProValIleSerLeu 346

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        347 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 366

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197
Db        367 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db        387 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
Db        407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
Db        427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
Db        447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT--- 1494
Db        467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 11
Q8BQY4
ID   Q8BQY4      PRELIMINARY;      PRT;   501 AA.
AC   Q8BQY4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
```

```
DE   Beta-site APP cleaving enzyme.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK046175; BAC32620.1; -.
SQ   SEQUENCE   501 AA;  55816 MW;  C0855513145E024E CRC64;

Alignment Scores:
Pred. No.:              7.07e-83        Length:         501
Score:                  1183.00         Matches:        237
Percent Similarity:     62.14%          Conservative:   83
Best Local Similarity:  46.02%          Mismatches:     169
Query Match:            35.43%          Indels:         26
DB:                     11              Gaps:           7

US-09-806-194-1 (1-1804) x Q8BQY4 (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla-------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCITGTCAACATTGCCACTATTTTTGAATCA 549
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCITATGCCACA 609
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGIGACACAAGCAAAC 669
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        247 GlyArgLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGIGTTTGATGCGGTGGTG 960
Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        307 LysSerIleLysAlaAlaSerSerThrGluLysProAspGlyPheTrpLeuGlyGlu 326

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIleSerLeu 346

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        347 IyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 366

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197
Db        367 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db        387 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
Db        407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
Db        427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
Db        447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCIGCTGCCGTTCCGGTGTCAGCGT--- 1494
Db        467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 12
Q8IYC8
ID   Q8IYC8      PRELIMINARY;      PRT;   501 AA.
AC   Q8IYC8;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Beta-site APP-cleaving enzyme.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Brain;
```

```
RA   Strausberg R.;
RL   Submitted (JUL-2002) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; BC036084; AAH36084.1; -.
SQ   SEQUENCE   501 AA;  55824 MW;  768595CF5517EFB7 CRC64;

Alignment Scores:
Pred. No.:              4.59e-82        Length:         501
Score:                  1172.50         Matches:        239
Percent Similarity:     61.85%          Conservative:   82
Best Local Similarity:  46.05%          Mismatches:     165
Query Match:            35.12%          Indels:         33
DB:                     4               Gaps:           9

US-09-806-194-1 (1-1804) x Q8IYC8 (1-501)

Qy         19 GCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC 78
Db          2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet---------GlyAlaGlyValLeuPro 18

Qy         79 GCGCCCTTCACG---------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC 123
Db         19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly--------- 35

Qy        124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183
Db         36 ----------------------GlyAlaProLeu--------------GlyLeuArg 42

Qy        184 CTCGCCCTGGAG------CCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCC 237
Db         43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62

Qy        238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG 297
Db         63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly 82

Qy        298 ACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA 357
Db         83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102

Qy        358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC 417
Db        103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuPheSerThrTyrArg 122

Qy        418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGG 477
Db        123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly 142

Qy        478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT 537
Db        143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162

Qy        538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGCATTAAATGGAATGGAATACTTGGCCTA 597
Db        163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182

Qy        598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG 657
Db        183 AlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal 202

Qy        658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT 717
Db        203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222

Qy        718 GCTGGATCT---------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db        223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis 242

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db        243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluVal 262

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db        283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysProAspGlyPhe 322

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068
Db        323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA 1185
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422

Qy       1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442

Qy       1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC 1425
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462

Qy       1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGG 1485
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg 481

Qy       1486 TGTCAGCGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 13
Q8C4F4
ID   Q8C4F4      PRELIMINARY;      PRT;   467 AA.
AC   Q8C4F4;
DT   01-MAR-2003 (TrEMBLrel. 23, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Beta-site APP cleaving enzyme.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
DR   EMBL; AK082317; BAC38462.1; -.
SQ   SEQUENCE   467 AA;  52063 MW;  31AB674FF1843652 CRC64;

Alignment Scores:
Pred. No.:              1.6e-72         Length:         467
Score:                  1049.00         Matches:        215
```

```
Percent Similarity:     56.70%          Conservative:   77
Best Local Similarity:  41.75%          Mismatches:     163
Query Match:            31.42%          Indels:         60
DB:                     11              Gaps:           8

US-09-806-194-1 (1-1804) x Q8C4F4 (1-467)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla-------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGlu------------------ 280

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        280 ------------------------------------------------------------ 280

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        281 ------------------------ThrGluLysPheProAspGlyPheTrpLeuGlyGlu 292

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        293 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIleSerLeu 312

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        313 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 332

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197
Db        333 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 352

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db        353 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 372

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
Db        373 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 392

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
Db        393 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 412

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
Db        413 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 432

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT--- 1494
Db        433 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 451

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        452 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 466
```

RESULT 14
Q9CUU5
ID   Q9CUU5   PRELIMINARY;   PRT;   267 AA.
AC   Q9CUU5;
DT   01-JUN-2001 (TrEMBLrel. 17, Created)
DT   01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT   01-MAR-2003 (TrEMBLrel. 23, Last annotation update)
DE   Adult male brain cDNA, RIKEN full-length enriched library,
DE   clone:3526402A15 product:beta-site APP cleaving enzyme, full insert
DE   sequence (Fragment).
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RA   Adachi J., Aizawa K., Akahira S., Akimura T., Arai A., Aono H.,
RA   Arakawa T., Bono H., Carninci P., Fukuda S., Fukunishi Y., Furuno M.,
RA   Hanagaki T., Hara A., Hayatsu N., Hiramoto K., Hiraoka T., Hori F.,
RA   Imotani K., Ishii Y., Itoh M., Izawa M., Kasukawa T., Kato H.,
RA   Kawai J., Kojima Y., Konno H., Kouda M., Koya S., Kurihara C.,
RA   Matsuyama T., Miyazaki A., Nishi K., Nomura K., Numazaki R., Ohno M.,
RA   Okazaki Y., Okido T., Owa C., Saito H., Saito R., Sakai C., Sakai K.,
RA   Sano H., Sasaki D., Shibata K., Shibata Y., Shinagawa A., Shiraki T.,
RA   Sogabe Y., Suzuki H., Tagami M., Tagawa A., Takahashi F., Tanaka T.,
RA   Tejima Y., Toya T., Yamamura T., Yasunishi A., Yoshida K., Yoshino M.,
RA   Muramatsu M., Hayashizaki Y.;
RL   Submitted (JUL-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=22354683; PubMed=12466851;
RA   The FANTOM Consortium,
RA   the RIKEN Genome Exploration Research Group Phase I & II Team;

```
RT   "Analysis of the mouse transcriptome based on functional annotation of
RT   60,770 full-length cDNAs.";
RL   Nature 420:563-573(2002).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=21085660; PubMed=11217851;
RA   RIKEN FANTOM Consortium;
RT   "Functional annotation of a full-length mouse cDNA collection.";
RL   Nature 409:685-690(2001).
RN   [4]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=99279253; PubMed=10349636;
RA   Carninci P., Hayashizaki Y.;
RT   "High-efficiency full-length cDNA cloning.";
RL   Meth. Enzymol. 303:19-44(1999).
RN   [5]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=20499374; PubMed=11042159;
RA   Carninci P., Shibata Y., Hayatsu N., Sugahara Y., Shibata K., Itoh M.,
RA   Konno H., Okazaki Y., Muramatsu M., Hayashizaki Y.;
RT   "Normalization and subtraction of cap-trapper-selected cDNAs to
RT   prepare full-length cDNA libraries for rapid discovery of new genes.";
RL   Genome Res. 10:1617-1630(2000).
RN   [6]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=20530913; PubMed=11076861;
RA   Shibata K., Itoh M., Aizawa K., Nagaoka S., Sasaki N., Carninci P.,
RA   Konno H., Akiyama J., Nishi K., Kitsunai T., Tashiro H., Itoh M.,
RA   Sumi N., Ishii Y., Nakamura S., Hazama M., Nishine T., Harada A.,
RA   Yamamoto R., Matsumoto H., Sakaguchi S., Ikegami T., Kashiwagi K.,
RA   Fujiwake S., Inoue K., Togawa Y., Izawa M., Ohara E., Watahiki M.,
RA   Yoneda Y., Ishikawa T., Ozawa K., Tanaka T., Matsuura S., Kawai J.,
RA   Okazaki Y., Muramatsu M., Inoue Y., Kira A., Hayashizaki Y.;
RT   "RIKEN integrated sequence analysis (RISA) system-384-format
RT   sequencing pipeline with 384 multicapillary sequencer.";
RL   Genome Res. 10:1757-1771(2000).
DR   EMBL; AK014390; BAB29317.2; -.
FT   NON_TER       1       1
SQ   SEQUENCE   267 AA;  30333 MW;  9413EB4530AB63B0 CRC64;

Alignment Scores:
Pred. No.:              5.68e-42        Length:         267
Score:                  653.00          Matches:        121
Percent Similarity:     66.29%          Conservative:   56
Best Local Similarity:  45.32%          Mismatches:     86
Query Match:            19.56%          Indels:         4
DB:                     11              Gaps:           3

US-09-806-194-1 (1-1804) x Q9CUU5 (1-267)

Qy        745 CTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATT 804
              :::::::::|||||||||:::   |||||||||   |||   :::|||||.|||||.|||
Db          1 MetIleIleGlyGlyIleAspHisSerLeuTyrThrGlySerLeuTrpTyrThrProIle 20

Qy        805 AAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTT 864
              :::   ||||||||||||||:::::   |||:::::::::|||||   ||||||   |||
Db         21 ArgArgGluTrpTyrTyrGluValIleIleValArgValGluIleAsnGlyGlnAspLeu 40

Qy        865 AATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTG 924
              :::|||||:::|||||||||   |||||||:::|||||||||||||||||||;|
Db         41 LysMetAspCysLysGluTyrAsnTyrAspLysSerIleValAspSerGlyThrThrAsn 60

Qy        925 CTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTG 984
              ||||||||||||:::|||||||||:::|||   |||:::::::::      ||||||
Db         61 LeuArgLeuProLysLysValPheGluAlaAlaValLysSerIleLysAlaAlaSerSer 80

Qy        985 ATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCG 1044
                 :::|||   ||||||||||||   |||   ||||||   ||||||
Db         81 ThrGluLysPheProAspGlyPheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGly 100

Qy       1045 GAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGC 1104
                 ||||||||||:::   ||||||   ||||||:::||||||      |||   ::::::
Db        101 ThrThrProTrpAsnIlePheProValIleSerLeuTyrLeuMetGlyGluValThrAsn 120

Qy       1105 AGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGC 1164
              :::||||||||||||||||||||||||||||   |||::::::|||:::
Db        121 GlnSerPheArgIleThrIleLeuProGlnGlnTyrLeuArgProValGluAspValAla 140

Qy       1165 CTGAATTAT---GAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATC 1221
                 :::       :::||||||:::|||   :::|||   |||:::         |||:::
Db        141 ThrSerGlnAspAspCysTyrLysPheAlaValSerGlnSerSerThrGlyThrValMet 160

Qy       1222 GGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCCAGAAGAGGGTGGGC 1281
              ||||||   :::||||||||||||||||||:::|||:||||||||:::|||||:::|||
Db        161 GlyAlaValIleMetGluGlyPheTyrValValPheAspArgAlaArgLysArgIleGly 180

Qy       1282 TTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCT 1341
              ||||||   |||   |||                          :::   :::   ||||||
Db        181 PheAlaValSerAlaCysHisValHisAspGluPheArgThrAlaAlaValGluGlyPro 200

Qy       1342 TTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATT 1401
              |||   |||   |||:::                                    |||
Db        201 PheValThrAlaAspMetGluAspCysGlyTyrAsnIleProGlnThrAspGluSerThr 220

Qy       1402 TTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATC 1461
              |||       :::::::|||   :::   :::::::|||   |||:::   ::::::   :::
Db        221 LeuMetThrIleAlaTyrValMetAlaAlaIleCys---AlaLeuPheMetLeuProLeu 239

Qy       1462 GTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT------CGCCCCCGTGACCCTGAGGTC 1515
              |||:::::::       :::||||||   |||      |||   :::      :::
Db        240 CysLeuMetValCysGlnTrpArgCysLeuArgCysLeuArgHisGlnHisAspAspPhe 259

Qy       1516 GTCAATGATGAGTCCTCTCTG 1536
              :::|||   |||   |||
Db        260 AlaAspAspIleSerLeuLeu 266

RESULT 15
Q9P0D2
ID   Q9P0D2      PRELIMINARY;      PRT;   213 AA.
AC   Q9P0D2;
DT   01-OCT-2000 (TrEMBLrel. 15, Created)
DT   01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT   01-JUN-2002 (TrEMBLrel. 21, Last annotation update)
DE   HSPC104 (Fragment).
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Umbilical cord blood;
RA   Zhang Q.H., Ye M., Zhou J., Shen Y., Wu X.Y., Guan Z.Q., Wang L.,
RA   Fan H.Y., Mao Y.F., Dai M., Huang Q.H., Chen S.J., Chen Z.;
RT   "Human partial CDS cloned from cd34+ stem cells.";
RL   Submitted (MAY-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF161367; AAF28927.1; -.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
FT   NON_TER       1       1
SQ   SEQUENCE   213 AA;  24338 MW;  EC9D3FA31CFA835C CRC64;

Alignment Scores:
Pred. No.:              3.75e-27        Length:         213
Score:                  461.00          Matches:        84
Percent Similarity:     68.10%          Conservative:   27
Best Local Similarity:  51.53%          Mismatches:     42
Query Match:            13.81%          Indels:         10
DB:                     4               Gaps:           1

US-09-806-194-1 (1-1804) x Q9P0D2 (1-213)

Qy        754 GGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAG 813
```

```
Db          1 GlyGlyIleAspHisSerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGlu 20

Qy        814 TGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGAC 873

Db         21 TrpTyrTyrGluValIleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAsp 40

Qy        874 TGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCCCTG 933

Db         41 CysLysGluTyrAsnTyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeu 60

Qy        934 CCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAA 993

Db         61 ProLysLysValPheGluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLys 80

Qy        994 TTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCT 1053

Db         81 PheProAspGlyPheTrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrPro 100

Qy       1054 TGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTC 1113

Db        101 TrpAsnIlePheProValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPhe 120

Qy       1114 CGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT 1173

Db        121 ArgIleThrIleLeuProGlnGlnTyrLeuArgPro---------------------- 132

Qy       1174 GAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTG 1233

Db        133 ------TrpLysMetTrpProArgProLysThrThrValThrValCysHisLeuThrVal 150

Qy       1234 ATGGAGGGC 1242

Db        151 IleHisGly 153


Search completed: September 25, 2003, 11:07:16
Job time : 204 secs
```

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:21:18 ; Search time 34.5 Seconds
(without alignments)
4918.039 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggcccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 127863 seqs, 47026705 residues

Total number of hits satisfying chosen parameters: 255726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095453_11005/app_query.fasta_1.1991
-DB=SwissProt_41 -QFMT=fastan -SUFFIX=rsp -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09806194_@CGN_1_1_46_@runat_25092003_095453_11005 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : SwissProt_41:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2687	80.5	518	1	BAE2_HUMAN	Q9y5z0 homo sapien
2	1187	35.5	501	1	BACE_RAT	P56819 rattus norv
3	1185	35.5	501	1	BACE_MOUSE	P56818 mus musculu
4	1178.5	35.3	501	1	BACE_HUMAN	P56817 homo sapien
5	363.5	10.9	377	1	PEPC_MACFU	P03955 macaca fusc
6	353	10.6	388	1	PEPC_HUMAN	P20142 homo sapien
7	351.5	10.5	388	1	PEPC_CALJA	Q9n2d3 callithrix
8	324.5	9.7	394	1	PEPC_CAVPO	Q64411 cavia porce
9	320	9.6	402	1	RENI_MOUSE	P06281 mus musculu
10	315.5	9.4	396	1	CATD_CLUHA	Q9dex3 clupea hare
11	313.5	9.4	509	1	APR1_ORYSA	Q42456 oryza sativ
12	313	9.4	392	1	PEPC_RAT	P04073 rattus norv
13	310	9.3	383	1	PEPE_CHICK	P16476 gallus gall
14	308.5	9.2	412	1	CATD_HUMAN	P07339 homo sapien
15	306.5	9.2	410	1	CATD_MOUSE	P18242 mus musculu
16	305.5	9.1	401	1	RENS_MOUSE	P00796 mus musculu
17	305	9.1	407	1	CATD_RAT	P24268 rattus norv
18	302	9.0	324	1	PEP1_GADMO	P56272 gadus morhu
19	302	9.0	405	1	CARP_YEAST	P07267 saccharomyc
20	301.5	9.0	398	1	CATE_RAT	P16228 rattus norv
21	300.5	9.0	387	1	PEP2_RABIT	P27821 oryctolagus
22	300.5	9.0	397	1	CATE_MOUSE	P70269 mus musculu
23	299	9.0	398	1	CATD_CHICK	Q05744 gallus gall
24	298.5	8.9	387	1	PEP4_RABIT	P28713 oryctolagus
25	298.5	8.9	391	1	CATE_CAVPO	P25796 cavia porce
26	298.5	8.9	400	1	RENI_SHEEP	P52115 ovis aries
27	297	8.9	388	1	PEPA_HUMAN	P00790 homo sapien
28	294.5	8.8	388	1	PEP2_MACFU	P27677 macaca fusc
29	291.5	8.7	388	1	PAG_HORSE	Q28389 equus cabal
30	291	8.7	388	1	PEP4_MACFU	P27678 macaca fusc
31	291	8.7	402	1	RENI_RAT	P08424 rattus norv
32	291	8.7	406	1	RENI_HUMAN	P00797 homo sapien
33	290.5	8.7	396	1	CATE_RABIT	P43159 oryctolagus
34	289	8.7	387	1	PEP3_RABIT	P27822 oryctolagus
35	288.5	8.6	390	1	CATD_BOVIN	P80209 bos taurus
36	288	8.6	387	1	PEP1_RABIT	P28712 oryctolagus
37	288	8.6	388	1	PEP1_MACFU	P03954 macaca fusc
38	287	8.6	367	1	PEPA_CHICK	P00793 gallus gall
39	287	8.6	396	1	CATE_HUMAN	P14091 homo sapien
40	286	8.6	388	1	PEPA_MACMU	P11489 macaca mula
41	285.5	8.6	387	1	PEPA_CALJA	Q9n2d4 callithrix
42	285	8.5	396	1	CARP_NEUCR	Q01294 neurospora
43	284.5	8.5	386	1	PEPA_PIG	P00791 sus scrofa
44	283	8.5	388	1	PEPF_RABIT	P27823 oryctolagus
45	282.5	8.5	381	1	CHYM_SHEEP	P18276 ovis aries

ALIGNMENTS

RESULT 1
BAE2_HUMAN
ID BAE2_HUMAN STANDARD; PRT; 518 AA.
AC Q9Y5Z0; Q9UJT6;
DT 30-MAY-2000 (Rel. 39, Created)
DT 30-MAY-2000 (Rel. 39, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Beta secretase 2 precursor (EC 3.4.23.-) (Beta-site APP-cleaving
DE enzyme 2) (Aspartyl protease 1) (Asp 1) (ASP1) (Membrane-associated
DE aspartic protease 1) (Memapsin-1).
GN BACE2 OR ASP21.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=20057170; PubMed=10591213;
RA Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT "Membrane-anchored aspartyl protease with Alzheimer's disease
RT beta-secretase activity.";
RL Nature 402:533-537(1999).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Bone marrow;
RA Xin H., Stephans J.C., Duan X., Harrowe G., Kim E., Grieshammer U.,
RA Giese K.;
RT "Identification of a novel aspartic-like protease differentially
RT expressed in human breast cancer cell lines.";
RL Submitted (JAN-1999) to the EMBL/GenBank/DDBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Accarino M.P., Fumagalli P., Ottolenghi S., Taramelli R.;
RT "Cloning of a gene from chromosome 21 Down region encoding a potential
RT transmembrane aspartyl protease.";
RL Submitted (FEB-1998) to the EMBL/GenBank/DDBJ databases.
RN [4]
RP SEQUENCE FROM N.A.
RA Solans A., Estivill X., de la Luna S.;
RT "Cloning of a novel mammalian aspartyl protease.";

```
RL   Submitted (AUG-1999) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20120043; PubMed=10656250;
RA   Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA   Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA   Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
RT   "Identification of a novel aspartic proteinase (Asp 2) as
RT   beta-secretase.";
RL   Mol. Cell. Neurosci. 14:419-427(1999).
RN   [6]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20144060; PubMed=10677483;
RA   Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT   "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
RT   beta-amyloid precursor protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RN   [7]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20289799; PubMed=10830953;
RA   Hattori M., Fujiyama A., Taylor T.D., Watanabe H., Yada T.,
RA   Park H.-S., Toyoda A., Ishii K., Totoki Y., Choi D.-K., Groner Y.,
RA   Soeda E., Ohki M., Takagi T., Sakaki Y., Taudien S., Blechschmidt K.,
RA   Polley A., Menzel U., Delabar J., Kumpf K., Lehmann R., Patterson D.,
RA   Reichwald K., Rump A., Schillhabel M., Schudy A., Zimmermann W.,
RA   Rosenthal A., Kudoh J., Shibuya K., Kawasaki K., Asakawa S.,
RA   Shintani A., Sasaki T., Nagamine K., Mitsuyama S., Antonarakis S.E.,
RA   Minoshima S., Shimizu N., Nordsiek G., Hornischer K., Brandt P.,
RA   Scharfe M., Schoen O., Desario A., Reichelt J., Kauer G., Bloecker H.,
RA   Ramser J., Beck A., Klages S., Hennig S., Riesselmann L., Dagand E.,
RA   Wehrmeyer S., Borzym K., Gardiner K., Nizetic D., Francis F.,
RA   Lehrach H., Reinhardt R., Yaspo M.-L.;
RT   "The DNA sequence of human chromosome 21.";
RL   Nature 405:311-319(2000).
RN   [8]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Skin;
RX   MEDLINE=22388257; PubMed=12477932;
RA   Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA   Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA   Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA   Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA   Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA   Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA   Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA   Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA   Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA   Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA   Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA   Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA   Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA   Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA   Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA   Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA   Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT   "Generation and initial analysis of more than 15,000 full-length
RT   human and mouse cDNA sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF200342; AAF17078.1; -.
DR   EMBL; AF117892; AAD45240.1; -.
DR   EMBL; AF050171; AAD45963.1; -.
DR   EMBL; AF178532; AAF29494.1; -.
DR   EMBL; AF204944; AAF26368.1; -.
DR   EMBL; AF200192; AAF13714.1; -.
DR   EMBL; AL163284; CAB90458.1; -.
DR   EMBL; AL163285; CAB90554.1; -.
DR   EMBL; BC014453; AAH14453.1; -.
DR   HSSP; P00797; 2REN.
DR   MEROPS; A01.041; -.
DR   Genew; HGNC:934; BACE2.
DR   MIM; 605668; -.
DR   GO; GO:0005624; C:membrane fraction; TAS.
DR   GO; GO:0006464; P:protein modification; TAS.
DR   GO; GO:0009306; P:protein secretion; TAS.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW   Signal.
FT   SIGNAL       1    20       POTENTIAL.
FT   PROPEP      21     ?       POTENTIAL.
FT   CHAIN        ?   518       BETA SECRETASE 2.
FT   DOMAIN      21   473       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM   474   494       POTENTIAL.
FT   DOMAIN     495   518       CYTOPLASMIC (POTENTIAL).
FT   ACT_SITE   110   110       BY SIMILARITY.
FT   ACT_SITE   303   303       BY SIMILARITY.
FT   CARBOHYD   170   170       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   366   366       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CONFLICT    36    36       A -> T (IN REF. 6).
SQ   SEQUENCE   518 AA;  56180 MW;  2E903150823760D3 CRC64;

Alignment Scores:
Pred. No.:                1.62e-171       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       1               Gaps:            0

US-09-806-194-1 (1-1804) x BAE2_HUMAN (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCICCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160
```

```
Qy       481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db       161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy       541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db       181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy       601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db       201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy       661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db       221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy       721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db       241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy       781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db       261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy       841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db       281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy       901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db       301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy       961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db       321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy      1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db       341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy      1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db       361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy      1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db       381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy      1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db       401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy      1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db       421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy      1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

```
RESULT 2
BACE_RAT
ID   BACE_RAT       STANDARD;      PRT:   501 AA.
AC   P56819;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE   (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE   protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
DE   (Memapsin-2).
GN   BACE.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20002972; PubMed=10531052;
RA   Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA   Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA   Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA   Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA   Treanor J., Rogers G., Citron M.;
RT   "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT   the transmembrane aspartic protease BACE.";
RL   Science 286:735-741(1999).
CC   -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC       AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC       THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
CC       LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC       SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC       FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF190727; AAF04144.1; -.
DR   HSSP; P32329; 1YPS.
DR   MEROPS; A01.004; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AsppproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     45       POTENTIAL.
FT   CHAIN        46    501       BETA-SECRETASE.
FT   DOMAIN       22    457       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458    478       POTENTIAL.
FT   DOMAIN      479    501       CYTOPLASMIC (POTENTIAL).
FT   ACT_SITE     93     93       BY SIMILARITY.
FT   ACT_SITE    289    289       BY SIMILARITY.
FT   DISULFID    216    420       BY SIMILARITY.
FT   DISULFID    278    443       BY SIMILARITY.
FT   DISULFID    330    380       BY SIMILARITY.
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    354    354       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   501 AA;  55806 MW;  24B445BC8BE87DE3 CRC64;

Alignment Scores:
Pred. No.:              1.32e-71       Length:        501
```

```
Score:                  1187.00         Matches:        240
Percent Similarity:     62.28%          Conservative:   82
Best Local Similarity:  46.42%          Mismatches:     165
Query Match:            35.55%          Indels:         30
DB:                     1               Gaps:           9

US-09-806-194-1 (1-1804) x BACE_RAT (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
              :::   |||   |||         ||||||                  :::   ||||||
Db          1 MetAlaProAlaLeuArgTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
                 |||              ||||||||||||      |||
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla-------------- 35

Qy        130 GCGCCCACCCCGGGACCC------GGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183
                  ||||||      |||          |||      |||
Db         36 ------------GlyProProLeuGlyLeuArgLeuProArgGluThrAspGlu------ 49

Qy        184 CTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTA 243
                 ||||||            |||            :::|||:::   |||||
Db         50 ---------GluPro------GluGluProGlyArgArgGlySerPheValGluMetVal 64

Qy        244 GACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCC 303
              ||||||||||:::|||   ||||||:::|||||||||:::|||||   :::|||:::||
Db         65 AspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerPro 84

Qy        304 CCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACC 363
              ||||||   |||   ||||||||||||||||||||:||||||:|||||||||||||||
Db         85 ProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAla 104

Qy        364 CCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAG 423
              ||||||   ::::::      |||:::      :::   ||||||||||||||||
Db        105 ProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeu 124

Qy        424 GGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGAC 483
                 |||   |||   |||||||||||:   ||:   |||   :::|||   |||
Db        125 ArgLysSerValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAsp 144

Qy        484 CTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTT 543
              ||||||:::||||||   |||   |||   :::          |:||||| |   |||
Db        145 LeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThr 164

Qy        544 GAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTAT 603
              |||||||:::   ||||||:::   |||      |||   ||||||||||||||:|||.
Db        165 GluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyr 184

Qy        604 GCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAA 663
              |||   :::|||:::|||      ||||||||   ||||||||||||||||||   |||
Db        185 AlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGln 204

Qy        664 GCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT------ 717
                 :::|||||||||:::||||||:::|||:::||||||||||||   |||:::
Db        205 ThrHisIleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGln 224

Qy        718 ---GCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTG 774
                    :::   |||||||||:::::::::|||||||||:::   |||||||
Db        225 ThrGluAlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeu 244

Cy        775 TATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATT 834
              |||   |||   :::|||||||||||||||||:::   ||||||||||:|::::::   |||
Db        245 TyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIle 264

Qy        835 CTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGAC 894
              :::::::::|||||   ||||||   |||   :::|||||||:::|||||||||   ||
Db        265 ValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAsp 284

Qy        895 AAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCG 954
              |||:::|||||||||||||||||||||   ||||||||||||:::|:|||||||:::|||
Db        285 LysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAla 304

Qy        955 GTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACT 1014
```

```
                  |||::::::::::      ||||||         :::|||   ||||||||||||
Db        305 AlaValLysSerIleLysAlaAlaSerSerThrGluLysProAspGlyPheTrpLeu 324

Qy       1015 GGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATC 1074
              |||   ||||||   ||||||            ||||||||||:::   ||||||   |||
Db        325 GlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIle 344

Qy       1075 TCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAG 1134
              |||:::||||||      |||   :::::::::||||||||||||||||||||||||||||
Db        345 SerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGln 364

Qy       1135 CTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGC 1191
                 |||::::::||||:::                  :::      :::|||||||:::|||
Db        365 GlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAla 384

Qy       1192 ATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTC 1251
              :::|||   |||:::            |||:::||||||   :::|||||||||||||||||
Db        385 ValSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrVal 404

Qy       1252 ATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCA 1311
              :::||||||||||||:::||||||:::||||||||   |||   |||
Db        405 ValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAsp 424

Qy       1312 GGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGT 1371
                    :::   :::   |||||||||   |||   |||:::
Db        425 GluPheArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGly 444

Qy       1372 GTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGC 1431
                                  |||      |||      ::::::|||   :::   :::
Db        445 TyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAla 464

Qy       1432 GTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAG 1491
              :::|||   |||:::   ::::::   :::   |||::::::      :::||||||
Db        465 IleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeu 483

Qy       1492 CGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
              |||      |||   :::      :::      :::|||   |||   |||
Db        484 ArgCysLeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500
```

```
RESULT 3
BACE_MOUSE
ID   BACE_MOUSE     STANDARD;      PRT;   501 AA.
AC   P56818;
DT   30-MAY-2000 (Rel. 39, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE   (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE   protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
DE   (Memapsin-2).
GN   BACE.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20002972; PubMed=10531052;
RA   Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA   Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA   Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA   Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA   Treanor J., Rogers G., Citron M.;
RT   "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT   the transmembrane aspartic protease BACE.";
RL   Science 286:735-741(1999).
RN   [2]
RP   REVISIONS TO 6 AND 81-87.
RA   Bennett B.D., Vassar R., Citron M.;
RL   Submitted (JAN-2000) to the EMBL/GenBank/DDBJ databases.
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=20057170; PubMed=10591213;
```

```
RA   Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA   Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA   Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT   "Membrane-anchored aspartyl protease with Alzheimer's disease
RT   beta-secretase activity.";
RL   Nature 402:533-537(1999).
CC   -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC       AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC       THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
CC       LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC       SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC       FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE (BY
CC       SIMILARITY).
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF190726; AAF04143.2; -.
DR   EMBL; AF200346; AAF17082.1; -.
DR   HSSP; P56272; 1AM5.
DR   MEROPS; A01.004; -.
DR   MGD; MGI:1346542; Bace.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW   Signal.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     45       POTENTIAL.
FT   CHAIN        46    501       BETA-SECRETASE.
FT   DOMAIN       22    457       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458    478       POTENTIAL.
FT   DOMAIN      479    501       CYTOPLASMIC (POTENTIAL).
FT   ACT_SITE     93     93       BY SIMILARITY.
FT   ACT_SITE    289    289       BY SIMILARITY.
FT   DISULFID    216    420       BY SIMILARITY.
FT   DISULFID    278    443       BY SIMILARITY.
FT   DISULFID    330    380       BY SIMILARITY.
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    354    354       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   501 AA;  55747 MW;  C085A013145E474E CRC64;

Alignment Scores:
Pred. No.:              1.79e-71      Length:          501
Score:                  1185.00       Matches:         237
Percent Similarity:     62.14%        Conservative:    83
Best Local Similarity:  46.02%        Mismatches:      169
Query Match:            35.49%        Indels:          26
DB:                     1             Gaps:            7

US-09-806-194-1 (1-1804) x BACE_MOUSE (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCCCC 84
              :::  |||  |||       ||||||             :::  ||||||
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
                 |||              ||||||||||||     |||
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla-------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
                 |||      |||         |||  |||         :::

Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249
                              |||            :::|||:::   ||||||||||||
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
              |||:::|||   ||||||:::||||||||||:::||||||   :::||||:::|||||||
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
                 |||   ||||||||||||||||||||||||||||||||||||||         ||||||
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
                 ::::::      ||:::      :::   ||||||||||||||
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
                 |||  |||  ||||||||||||  |||  |||  :::|||  |||||||||
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
              :::||||||   |||   |||  :::          |||||||||   |||  ||||||
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609
              :::   |||||||:::   |||      |||   ||||||||||||||||||||||||
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669
              :::|||:::|||      |||||||||  ||||||||||||||||||  |||  :::
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
              |||||||||:::||||||:::|||:::||||||||||||   |||:::
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
                    :::   |||||||||:::::::::|||||||||:::   ||||||||
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              |||   :::|||||:|||||||||:::   ||||||||||||:::::   |||::::::
Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              :::|||||||   ||||||   |||   :::|||||||:::||||||||   ||||||:::
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||   |||||:||||||:::|||||||||:::|||   |||
Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              :::::::::      ||||||          :::|||   :||||||||||   |||
Db        307 LysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGlu 326

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||   ||||||              ||||||||:::   ||||||   ||||||:::
Db        327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIleSerLeu 346

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||      |||  :::::::::||||||||||||||||||||||||||||   |||
Db        347 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 366

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197
              ::::::|||:::              :::      :::|||||||:::|||   :::|||
Db        367 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
                 |||:::        |||:::||||||   :::|||||||||||||||||:::|||
Db        387 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406
```

```
Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
             ||||||||||:::||||||:::||||||||    |||   |||
Db        407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
                 :::   :::   ||||||||   |||   ||:::
Db        427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
                    |||      |||      ::::::|||   :::   :::::|||
Db        447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT--- 1494
                |||:::   ::::::   :::   |||::::::      :::|||||   |||
Db        467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
                |||   :::      :::      :::|||   |||   |||
Db        486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 4
BACE_HUMAN
ID   BACE_HUMAN     STANDARD;      PRT;   501 AA.
AC   P56817; Q9BYB9; Q9BYC0; Q9BYC1; Q9UJT5;
DT   30-MAY-2000 (Rel. 39, Created)
DT   30-MAY-2000 (Rel. 39, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Beta-secretase precursor (EC 3.4.23.-) (Beta-site APP cleaving enzyme)
DE   (Beta-site amyloid precursor protein cleaving enzyme) (Aspartyl
DE   protease 2) (Asp 2) (ASP2) (Membrane-associated aspartic protease 2)
DE   (Memapsin-2).
GN   BACE OR BACE1.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RC   TISSUE=Brain;
RX   MEDLINE=20002972; PubMed=10531052;
RA   Vassar R., Bennett B.D., Babu-Khan S., Kahn S., Mendiaz E.A.,
RA   Denis P., Teplow D.B., Ross S., Amarante P., Loeloff R., Luo Y.,
RA   Fisher S., Fuller J., Edenson S., Lile J., Jarosinski M.A.,
RA   Biere A.L., Curran E., Burgess T., Louis J.-C., Collins F.,
RA   Treanor J., Rogers G., Citron M.;
RT   "Beta-secretase cleavage of Alzheimer's amyloid precursor protein by
RT   the transmembrane aspartic protease BACE.";
RL   Science 286:735-741(1999).
RN   [2]
RP   SEQUENCE FROM N.A. (ISOFORM A), SEQUENCE OF 46-68, AND
RP   CHARACTERIZATION.
RC   TISSUE=Brain;
RX   MEDLINE=20057171; PubMed=10591214;
RA   Sinha S., Anderson J.P., Barbour R., Basi G.S., Caccavello R.,
RA   Davis D., Doan M., Dovey H.F., Frigon N., Hong J., Jacobson-Croak K.,
RA   Jewett N., Keim P., Knops J., Lieberburg I., Power M., Tan H.,
RA   Tatsuno G., Tung J., Schenk D., Seubert P., Suomensaari S.M., Wang S.,
RA   Walker D., Zhao J., McConlogue L., Varghese J.;
RT   "Purification and cloning of amyloid precursor protein beta-secretase
RT   from human brain.";
RL   Nature 402:537-540(1999).
RN   [3]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=20057170; PubMed=10591213;
RA   Yan R., Bienkowski M.J., Shuck M.E., Miao H., Tory M.C., Pauley A.M.,
RA   Brashier J.R., Stratman N.C., Mathews W.R., Buhl A.E., Carter D.B.,
RA   Tomasselli A.G., Parodi L.A., Heinrikson R.L., Gurney M.E.;
RT   "Membrane-anchored aspartyl protease with Alzheimer's disease beta-
RT   secretase activity.";
RL   Nature 402:533-537(1999).
RN   [4]
RP   SEQUENCE FROM N.A. (ISOFORM A).
RX   MEDLINE=20120043; PubMed=10656250;
RA   Hussain I., Powell D.J., Howlett D.R., Tew D.G., Meek T.D.,
RA   Chapman C., Gloger I.S., Murphy K.E., Southan C.D., Ryan D.M.,
RA   Smith T.S., Simmons D.L., Walsh F.S., Dingwall C., Christie G.;
RT   "Identification of a novel aspartic proteinase (Asp 2) as beta-
RT   secretase.";
RL   Mol. Cell. Neurosci. 14:419-427(1999).
RN   [5]
RP   SEQUENCE FROM N.A. (ISOFORM B).
RC   TISSUE=Brain, and Pancreas;
RA   Michel B., De Pietri Tonelli D., Zacchetti D., Keller P.;
RT   "New beta-site APP cleaving enzyme isoform (BACE-1B) obtained from
RT   human brain and pancreas.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [6]
RP   SEQUENCE FROM N.A. (ISOFORM C).
RC   TISSUE=Pancreas;
RA   Zacchetti D., De Pietri Tonelli D., Schnurbus R.;
RT   "New beta-site APP cleaving enzyme isoform (BACE-1C) obtained from
RT   human pancreas.";
RL   Submitted (JAN-2001) to the EMBL/GenBank/DDBJ databases.
RN   [7]
RP   SEQUENCE FROM N.A. (ISOFORMS B; C AND D).
RC   TISSUE=Brain;
RX   MEDLINE=21408467; PubMed=11516562;
RA   Tanahashi H., Tabira T.;
RT   "Three novel alternatively spliced isoforms of the human beta-site
RT   amyloid precursor protein cleaving enzyme (BACE) and their effect on
RT   amyloid beta-peptide production.";
RL   Neurosci. Lett. 307:9-12(2001).
RN   [8]
RP   SEQUENCE OF 14-501 FROM N.A. (ISOFORM A), AND CHARACTERIZATION.
RX   MEDLINE=20144060; PubMed=10677483;
RA   Lin X., Koelsch G., Wu S., Downs D., Dashti A., Tang J.;
RT   "Human aspartic protease memapsin 2 cleaves the beta-secretase site of
RT   beta-amyloid precursor protein.";
RL   Proc. Natl. Acad. Sci. U.S.A. 97:1456-1460(2000).
RN   [9]
RP   DISULFIDE BONDS.
RX   MEDLINE=21950860; PubMed=11953458;
RA   Fischer F., Molinari M., Bodendorf U., Paganetti P.;
RT   "The disulphide bonds in the catalytic domain of BACE are critical but
RT   not essential for amyloid precursor protein processing activity.";
RL   J. Neurochem. 80:1079-1088(2002).
CC   -!- FUNCTION: RESPONSIBLE FOR THE PROTEOLYTIC PROCESSING OF THE
CC       AMYLOID PRECURSOR PROTEIN (APP). CLEAVES AT THE AMINO TERMINUS OF
CC       THE A-BETA PEPTIDE SEQUENCE, BETWEEN RESIDUES 671 AND 672 OF APP,
CC       LEADS TO THE GENERATION AND EXTRACELLULAR RELEASE OF BETA-CLEAVED
CC       SOLUBLE APP, AND A CORRESPONDING CELL-ASSOCIATED CARBOXY-TERMINAL
CC       FRAGMENT WHICH IS LATER RELEASE BY GAMMA-SECRETASE.
CC   -!- SUBCELLULAR LOCATION: Type I membrane protein.
CC   -!- ALTERNATIVE PRODUCTS:
CC       Event=Alternative splicing; Named isoforms=4;
CC       Name=A; Synonyms=BACE-1A, BAC-501;
CC         IsoId=P56817-1; Sequence=Displayed;
CC       Name=B; Synonyms=BACE-1B, BACE-I-476;
CC         IsoId=P56817-2; Sequence=VSP_005223;
CC       Name=C; Synonyms=BACE-1C, BACE-I-457;
CC         IsoId=P56817-3; Sequence=VSP_005222;
CC       Name=D; Synonyms=BACE-1D, BACE-I-432;
CC         IsoId=P56817-4; Sequence=VSP_005222, VSP_005223;
CC   -!- TISSUE SPECIFICITY: BRAIN.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF190725; AAF04142.1; -.
```

```
DR   EMBL; AF201468; AAF18982.1; -.
DR   EMBL; AF200343; AAF17079.1; -.
DR   EMBL; AF204943; AAF26367.1; -.
DR   EMBL; AF338816; AAK38374.1; -.
DR   EMBL; AF338817; AAK38375.1; -.
DR   EMBL; AB050436; BAB40931.1; -.
DR   EMBL; AB050437; BAB40932.1; -.
DR   EMBL; AB050438; BAB40933.1; -.
DR   EMBL; AF200193; AAF13715.1; -.
DR   PIR; A59090; A59090.
DR   PDB; 1M4H; 28-AUG-02.
DR   MEROPS; A01.004; -.
DR   Genew; HGNC:933; BACE.
DR   MIM; 604252; -.
DR   GO; GO:0005887; C:integral to plasma membrane; TAS.
DR   GO; GO:0008798; F:beta-aspartyl-peptidase activity; TAS.
DR   GO; GO:0009405; P:pathogenesis; TAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AsppproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 1.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Zymogen; Transmembrane;
KW   Signal; Alternative splicing; 3D-structure.
FT   SIGNAL        1     21       POTENTIAL.
FT   PROPEP       22     45
FT   CHAIN        46    501       BETA-SECRETASE.
FT   DOMAIN       22    457       EXTRACELLULAR (POTENTIAL).
FT   TRANSMEM    458    478       POTENTIAL.
FT   DOMAIN      479    501       CYTOPLASMIC (POTENTIAL).
FT   ACT_SITE     93     93       BY SIMILARITY.
FT   ACT_SITE    289    289       BY SIMILARITY.
FT   DISULFID    216    420
FT   DISULFID    278    443
FT   DISULFID    330    380
FT   CARBOHYD    153    153       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    172    172       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    223    223       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    354    354       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARSPLIC    146    189       Missing (in isoform C and isoform D).
FT                                /FTId=VSP_005222.
FT   VARSPLIC    190    214       Missing (in isoform B and isoform D).
FT                                /FTId=VSP_005223.
SQ   SEQUENCE   501 AA;  55763 MW;  377CE4C824ACEF05 CRC64;

Alignment Scores:
Pred. No.:              4.85e-71      Length:         501
Score:                  1178.50       Matches:        240
Percent Similarity:     62.04%        Conservative:   82
Best Local Similarity:  46.24%        Mismatches:     164
Query Match:            35.29%        Indels:         33
DB:                     1             Gaps:           9

US-09-806-194-1 (1-1804) x BACE_HUMAN (1-501)

Qy         19 GCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC 78
              |||     ||||||   |||     |||:::           :::   |||
Db          2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet---------GlyAlaGlyValLeuPro 18

Qy         79 GCGCCCTTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC 123
              |||     |||              ||||||||||
Db         19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly--------- 35

Qy        124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183
                                  |||   |||             ||||||
Db         36 ----------------------GlyAlaProLeu--------------GlyLeuArg 42

Qy        184 CTCGCCCTGGAG------CCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCC 237
              |||     |||                 |||          :::|||:::
Db         43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62

Qy        238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG 297
              |||||||||||||||:::|||   ||||||:::|||||||||:::||||||   :::|||
Db         63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly 82

Qy        298 ACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA 357
              :::|||||||||   |||   |||||||||||||||||||||||||||||||||||||
Db         83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102

Qy        358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC 417
                 ||||||   ::::::      |||:::      :::   ||||||||||||||||
Db        103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg 122

Qy        418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGG 477
                       |||   |||   ||||||||||||   |||   |||   :::|||
Db        123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly 142

Qy        478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT 537
                 ||||||||||:::||||||   |||   |||   :::          ||||||||
Db        143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162

Qy        538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTA 597
                 |||   |||||||:::   |||||||:::   |||       |||   ||||||||||||||
Db        163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182

Qy        598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG 657
                 |||||||||   :::|||:::|||       |||||||||   ||||||||||||||||||
Db        183 AlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal 202

Qy        658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT 717
                 |||   ::::::|||||||:::||||||:::|||:::||||||||||   |||:::
Db        203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222

Qy        718 GCTGGATCT---------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
                 |||             :::   |||||||||:::::::::|||||||||:::
Db        223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis 242

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
                 |||||||||   ||;   :::|||||||||||||||||:::   |||||||||||::::::
Db        243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluVal 262

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
                 |||:::::::::||||||   ||||||   |||   :::||||||:::|||||||||
Db        263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
                 ||||||:::|||||||||||||||||||   ||||||||||||:::|||||||||
Db        283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302

Qy        949 GATGCCGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
                 :::|||   |||:::::::::       ||||||          :::|||   |||||||||
Db        303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe 322

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068
                 |||   |||   ||||||   ||||||            |||||||||:::   ||||||
Db        323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
                 ||||||:::||||||       |||   :::::::::|||||||||||||||||||||||
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA 1185
                 ||||||   |||:::::::|||:::           :::       :::||||||||:::
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
                 |||   ||||||   |||:::          |||:::||||||   :::|||||||||||
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
                 ||||||:::|||||||||||||:::||||||:::||||||||||   |||   |||
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422

Qy       1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
                          :::   :::   |||||||||   |||   |||:::
```

```
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442

Qy       1366 AACTGTGTCCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC 1425
                                     |||      |||      :::::::||   :::
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462

Qy       1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGG 1485
              :::::::|||   |||:::   ::::::   :::   |||:::::::      :::|||
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg 481

Qy       1486 TGTCAGCGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
              |||   |||      |||   :::      :::      :::|||   |||   |||
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 5
PEPC_MACFU
ID   PEPC_MACFU     STANDARD;      PRT;   377 AA.
AC   P03955;
DT   23-OCT-1986 (Rel. 02, Created)
DT   01-AUG-1992 (Rel. 23, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C) (Fragment).
GN   PGC.
OS   Macaca fuscata fuscata (Japanese macaque).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Cercopithecidae;
OC   Cercopithecinae; Macaca.
OX   NCBI_TaxID=9543;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Gastric mucosa;
RX   MEDLINE=92037645; PubMed=1935977;
RA   Kageyama T., Tanabe K., Koiwai O.;
RT   "Development-dependent expression of isozymogens of monkey
RT   pepsinogens and structural differences between them.";
RL   Eur. J. Biochem. 202:205-215(1991).
RN   [2]
RP   SEQUENCE OF 6-377.
RX   MEDLINE=86168133; PubMed=3514597;
RA   Kageyama T., Takahashi K.;
RT   "The complete amino acid sequence of monkey progastricsin.";
RL   J. Biol. Chem. 261:4406-4419(1986).
RN   [3]
RP   SEQUENCE OF 6-65.
RX   MEDLINE=85289106; PubMed=3928607;
RA   Kageyama T., Takahashi K.;
RT   "Monkey pepsinogens and pepsins. VII. Analysis of the activation
RT   process and determination of the NH2-terminal 60-residue sequence of
RT   Japanese monkey progastricsin, and molecular evolution of
RT   pepsinogens.";
RL   J. Biochem. 97:1235-1246(1985).
CC   -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC       shows preferential cleavage at Tyr-|-Xaa bonds; high activity
CC       towards hemoglobin as substrate.
CC   -!- PTM: EACH PEPSINOGEN IS CONVERTED TO CORRESPONDING PEPSIN AT PH
CC       2.0 IN PART AS A RESULT OF THE RELEASE OF A 47 AA ACTIVATION
CC       SEGMENT AND IN PART AS A RESULT OF STEPWISE PROTEOLYTIC CLEAVAGE
CC       VIA AN INTERMEDIATE FORM(S).
CC   -!- MISCELLANEOUS: THE EXPRESSION OF PEPSINOGEN GENES IS REGULATED BY
CC       HORMONES AND RELATED SUBSTANCES.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X59754; CAA42426.1; -.
DR   PIR; S19683; PEMQCJ.
DR   HSSP; P20142; 1AVF.
DR   MEROPS; A01.003; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
FT   NON_TER       1      1
FT   SIGNAL       <1      5
FT   PROPEP        6     31       ACTIVATION PEPTIDE.
FT   PROPEP       32     48       ACTIVATION PEPTIDE.
FT   CHAIN        49    377       GASTRICSIN.
FT   DISULFID     93     98
FT   DISULFID    256    260
FT   DISULFID    299    332
FT   ACT_SITE     80     80
FT   ACT_SITE    265    265
FT   CONFLICT    331    331       Y -> V (IN REF. 2).
FT   CONFLICT    349    349       L -> LVY (IN REF. 2).
SQ   SEQUENCE   377 AA;  41148 MW;  2CFB8F8BF26D77CE CRC64;

Alignment Scores:
Pred. No.:              8.87e-17        Length:         377
Score:                  363.50          Matches:        118
Percent Similarity:     44.85%          Conservative:   65
Best Local Similarity:  28.92%          Mismatches:     118
Query Match:            10.89%          Indels:         107
DB:                     1               Gaps:           19

US-09-806-194-1 (1-1804) x PEPC_MACFU (1-377)

Qy        166 CACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCC 225
              |||          |||::::::::::   ||||||
Db         44 HisPheGlyAspLeuSerValSerTyrGluPro-------------------------- 54

Qy        226 AACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAG 285
                  ::::|||   :::|||                                |||:::   |||
Db         55 ------MetAlaTyrMetAsp--------------------AlaAlaTyrPheGlyGlu 65

Qy        286 ATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC 345
                  :::   ||||||||||||||||||          :::|||   ||||||||||||||||||
Db         66 IleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsn 85

Qy        346 TTTGCCGTG----------------------GCAGGAACCCCGCACTCCTACATAGAC 381
                  |||                        |||   |||   ||||||
Db         86 LeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHisSer--------- 102

Qy        382 ACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTG 441
                  |||:::          ||||||||||   :::   |||         ::::::
Db        103 ---ArgPheAsnProSerGluSerSerThrTyrSerThrAsnGlyGlnThrPheSerLeu 121

Qy        442 AAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAA 501
                  :::|||      ||||||   |||||||||   |||   |||   :::|||:::
Db        122 GlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeuThrVal------ 139

Qy        502 GGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTT 561
                       |||      |||                        |||||||||
Db        140 ---------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSerGluAsn------ 154

Qy        562 TTGCCTGGG--------------ATTAAATGGAATGGAATACTTGGCCTAGCTTATGCC 606
                  ||||||                  :::::::::||||||:::||||||||||||
Db        155 GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGlyLeuAlaTyrPro 174

Qy        607 ACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCA 666
                  ||||||:::        ::::::     |||          :::|||   :::
Db        175 ThrLeuSerValAspGlyAlaThr------ThrAlaMetGlnGlyMetValGlnGluGly 192

Qy        667 AACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGA 723
                  :::   :::   :::||||||:::   :::
Db        193 AlaLeuThrSerProIlePheSerValTyrLeuSerAspGln------------------ 206

Qy        724 TCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGA 783
                  |||::::::|||||||:::::|||   ||||||:::::   |||||||||   |||
```

```
Db        207 GlnGlySerSerGlyGlyAlaValValPheGlyGlyValAspSerSerLeuTyrThrGly 226

Qy        784 GACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTG 843

Db        227 GlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGlyIleGluPhe 246

Qy        844 GAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATC 903

Db        247 LeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------GlyCysGlnAlaIle 283

Qy        904 GTGGACAGTGGCACCACGCTGCTGCCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAA 963

Db        264 ValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSerAlaLeuLeuGln 283

Qy        964 GCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG 1023

Db        284 Ala----------------------------------------------ThrGlyAlaGln 288

Qy       1024 CTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC---------------- 1065

Db        289 --------------------GluAspGluTyrGlyGlnPheLeuValAsnCysAsnSer 301

Qy       1066 ------------CCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTC 1113

Db        302 IleGlnAsnLeuProThrLeuThrPheIleIle------------------AsnGlyVal 315

Qy       1114 CGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT 1173

Db        316 GluPheProLeuProProSerSerTyrIle--------------------LeuAsnAsn 328

Qy       1174 GAATGTTAC---CGATTCGGCATTTCCCCA----------------------TCCACAAAT 1209

Db        329 AsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerAlaGlnAsnSerGlnPro 348

Qy       1210 GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAG 1269

Db        349 LeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuSerAsn 368

Qy       1270 AAGAGGGTGGGCTTCGCAGCGAGC 1293

Db        369 AsnArgValGlyPheAlaThrAla 376

RESULT 6
PEPC_HUMAN
ID   PEPC_HUMAN     STANDARD;      PRT;   388 AA.
AC   P20142;
DT   01-FEB-1991 (Rel. 17, Created)
DT   01-FEB-1991 (Rel. 17, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
GN   PGC.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88087276; PubMed=3335549;
RA   Hayano T., Sogawa K., Ichihara Y., Fujii-Kuriyama Y., Takahashi K.;
RT   "Primary structure of human pepsinogen C gene.";
RL   J. Biol. Chem. 263:1382-1385(1988).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=89079679; PubMed=2909526;
RA   Taggart R.T., Cass L.G., Mohandas T.K., Derby P., Barr P.J., Pals G.,
RA   Bell G.I.;
RT   "Human pepsinogen C (progastricsin). Isolation of cDNA clones,
RT   localization to chromosome 6, and sequence homology with pepsinogen
RT   A.";
RL   J. Biol. Chem. 264:375-379(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   TISSUE=Placenta;
RX   MEDLINE=89290840; PubMed=2567697;
RA   Pals G., Azuma T., Mohandas T.K., Bell G.I., Bacon J.,
RA   Samloff I.M., Walz D.A., Barr P.J., Taggart R.T.;
RT   "Human pepsinogen C (progastricsin) polymorphism: evidence for a
RT   single locus located at 6p21.1-pter.";
RL   Genomics 4:137-148(1989).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Wong R.N.S., Tang J.;
RL   Submitted (NOV-1996) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE OF 17-101.
RX   MEDLINE=90130402; PubMed=2515193;
RA   Athauda S.B.P., Tanji M., Kageyama T., Takahashi K.;
RT   "A comparative study on the NH2-terminal amino acid sequences and
RT   some other properties of six isozymic forms of human pepsinogens and
RT   pepsins.";
RL   J. Biochem. 106:920-927(1989).
RN   [6]
RP   SEQUENCE OF 17-64.
RX   MEDLINE=83079318; PubMed=6816595;
RA   Foltmann B., Jensen A.L.;
RT   "Human progastricsin. Analysis of intermediates during activation
RT   into gastricsin and determination of the amino acid sequence of the
RT   propart.";
RL   Eur. J. Biochem. 128:63-70(1982).
RN   [7]
RP   X-RAY CRYSTALLOGRAPHY (1.62 ANGSTROMS).
RX   MEDLINE=95230687; PubMed=7714902;
RA   Moore S.A., Sielecki A.R., Chernaia M.M., Tarasova N.I., James M.N.G.;
RT   "Crystal and molecular structures of human progastricsin at 1.62-A
RT   resolution.";
RL   J. Mol. Biol. 247:466-485(1995).
RN   [8]
RP   X-RAY CRYSTALLOGRAPHY (2.36 ANGSTROMS).
RX   MEDLINE=98069649; PubMed=9406551;
RA   Khan A.R., Cherney M.M., Tarasova N.I., James M.N.;
RT   "Structural characterization of activation 'intermediate 2' on the
RT   pathway to human gastricsin.";
RL   Nat. Struct. Biol. 4:1010-1015(1997).
CC   -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC       shows preferential cleavage at Tyr-|-Xaa bonds; high activity
CC       towards hemoglobin as substrate.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M18667; AAA60062.1; ALT_INIT.
DR   EMBL; M18659; AAA60062.1; JOINED.
DR   EMBL; M18660; AAA60062.1; JOINED.
DR   EMBL; M18661; AAA60062.1; JOINED.
DR   EMBL; M18662; AAA60062.1; JOINED.
DR   EMBL; M18663; AAA60062.1; JOINED.
DR   EMBL; M18664; AAA60062.1; JOINED.
DR   EMBL; M18665; AAA60062.1; JOINED.
DR   EMBL; M18666; AAA60062.1; JOINED.
DR   EMBL; M23077; AAA60063.1; -.
DR   EMBL; M23069; AAA60063.1; JOINED.
DR   EMBL; M23070; AAA60063.1; JOINED.
DR   EMBL; M23071; AAA60063.1; JOINED.
DR   EMBL; M23072; AAA60063.1; JOINED.
DR   EMBL; M23073; AAA60063.1; JOINED.
DR   EMBL; M23074; AAA60063.1; JOINED.
DR   EMBL; M23075; AAA60063.1; JOINED.
DR   EMBL; J04443; AAA60074.1; -.
DR   EMBL; U75272; AAB18273.1; -.
DR   PIR; A29937; A29937.
```

```
DR   PDB; 1HTR; 26-JAN-95.
DR   PDB; 1AVF; 25-FEB-98.
DR   MEROPS; A01.003; -.
DR   Genew; HGNC:8890; PGC.
DR   MIM; 169740; -.
DR   GO; GO:0005615; C:extracellular space; TAS.
DR   GO; GO:0007586; P:digestion; TAS.
DR   GO; GO:0006508; P:proteolysis and peptidolysis; TAS.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal;
KW   3D-structure.
FT   SIGNAL       1    16
FT   PROPEP      17    59       ACTIVATION PEPTIDE.
FT   CHAIN       60   388       GASTRICSIN.
FT   ACT_SITE    91    91
FT   ACT_SITE   276   276
FT   DISULFID   104   109
FT   DISULFID   267   271
FT   DISULFID   310   343
FT   CONFLICT    40    41       GE -> ED (IN REF. 6).
FT   CONFLICT    52    52       W -> S (IN REF. 6).
FT   STRAND      19    25
FT   HELIX       29    35
FT   TURN        36    37
FT   HELIX       39    43
FT   TURN        44    45
FT   HELIX       50    54
FT   HELIX       65    68
FT   TURN        69    70
FT   STRAND      73    79
FT   TURN        80    83
FT   STRAND      84    91
FT   TURN        92    93
FT   STRAND      97   101
FT   TURN       102   103
FT   HELIX      107   110
FT   TURN       111   111
FT   STRAND     115   115
FT   HELIX      117   119
FT   TURN       121   122
FT   STRAND     124   134
FT   TURN       135   136
FT   STRAND     137   150
FT   TURN       151   152
FT   STRAND     153   163
FT   HELIX      169   173
FT   STRAND     178   181
FT   TURN       190   191
FT   HELIX      195   201
FT   TURN       202   203
FT   STRAND     209   214
FT   STRAND     221   227
FT   HELIX      232   234
FT   STRAND     235   244
FT   STRAND     251   254
FT   STRAND     256   259
FT   TURN       260   261
FT   STRAND     262   263
FT   TURN       266   269
FT   STRAND     271   275
FT   TURN       277   278
FT   STRAND     282   285
FT   HELIX      286   288
FT   HELIX      289   296
FT   TURN       297   297
FT   STRAND     299   300
FT   TURN       302   303
FT   STRAND     306   308
FT   HELIX      310   315
FT   STRAND     319   323
FT   TURN       324   325
FT   STRAND     326   330
FT   HELIX      332   335
FT   STRAND     336   338
FT   STRAND     343   345
FT   STRAND     347   350
FT   TURN       355   356
FT   STRAND     360   363
FT   HELIX      365   368
FT   TURN       369   370
FT   STRAND     371   376
FT   TURN       377   380
FT   STRAND     381   388
SQ   SEQUENCE   388 AA;  42426 MW;  F862DFDC1438BB92 CRC64;

Alignment Scores:
Pred. No.:                4.46e-16        Length:          388
Score:                    353.00          Matches:         120
Percent Similarity:       44.79%          Conservative:    65
Best Local Similarity:    29.06%          Mismatches:      120
Query Match:              10.57%          Indels:          108
DB:                       1               Gaps:            21

US-09-806-194-1 (1-1804) x PEPC_HUMAN (1-388)

Qy        154 CCTGCCGAGCGCCACGCCGACGGC---TTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCC 210
              ||||||   ::::::      |||   |||::::::      ||||||
Db         50 ProAlaTrpLysTyrArgPheGlyAspLeuSerValThrTyrGluPro------------ 65

Qy        211 CCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGC 270
                                  :::|||   :::|||
Db         66 --------------------MetAlaTyrMetAsp--------------------Ala 71

Qy        271 GGCTACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGAC 330
                 |||:::   |||:::   ||||||||||||||||!          :::|||   |||
Db         72 AlaTyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAsp 91

Qy        331 ACTGGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCG 366
              ;|||||||||||||||      |||                     |||   |||
Db         92 ThrGlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSer 111

Qy        367 CACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGC 426
              ||||||               |||:::          ||||||||||||   :::   |||
Db        112 HisSer------------ArgPheAsnProSerGluSerSerThrTyrSerThrAsnGly 127

Qy        427 TTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTC 486
                  ::::::::::|||      ||||||   ||||||||   |||   |||
Db        128 GlnThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThr 147

Qy        487 GTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAA 546
              :::|||:::           |||   |||
Db        148 LeuThrVal--------------GlnSerIleGlnValProAsnGlnGluPheGlyLeu 162

Qy        547 TCAGAGAATTTCTTTTTGCCTGGG--------------ATTAAATGGAATGGAATACTT 591
              |||||||||         ||||||                  :::::::::|||||:::
Db        163 SerGluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMet 180

Qy        592 GGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC 651
              ||||||||||||      |||:::          ::::::      |||
Db        181 GlyLeuAlaTyrProAlaLeuSerValAspGluAlaThr------ThrAlaMetGlnGly 198

Qy        652 CTGGTGACACAAGCAAACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGC 708
              :::|||   :::       :::   :::   ||||||||:::   :::
Db        199 MetValGlnGluGlyAlaLeuThrSerProValPheSerValTyrLeuSerAsnGln--- 217

Qy        709 TTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
                          |||::::::|||||:::::|||   ||||||::::::
Db        218 --------------GlnGlySerSerGlyGlyAlaValValPheGlyGlyValAspSer 232

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
              |||||||||   |||   |||:::::::   |||:::   :::|||   |||:::||||||
Db        233 SerLeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIle 252
```

```
Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        253 GlyIleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------ 289

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCCCTGCCCCAGAAGGTGTTT 948
Db        270 GlyCysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMet 289

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        290 SerAlaLeuLeuGlnAla---------------------------------------- 295

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC--- 1065
Db        296 ---ThrGlyAlaGln--------------------GluAspGluTyrGlyGlnPheLeu 307

Qy       1066 --------------------------CCTAAAATCTCCATCTACCTGAGAGATGAGAAC 1098
Db        308 ValAsnCysAsnSerIleGlnAsnLeuProSerLeuThrPheIleIle------------ 323

Qy       1099 TCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGG 1158
Db        324 ------AsnGlyValGluPheProLeuProProSerSerTyrIle--------------- 336

Qy       1159 GCCGGCCTGAATTATGAATGTTAC---CGATTCGGCATTTCCCCA------------TCC 1203
Db        337 ------LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSer 354

Qy       1204 ACAAATGCG---------CTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATC 1254
Db        355 GlnAsnGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerVal 374

Qy       1255 TTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
Db        375 TyrAspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387
```

RESULT 7
PEPC_CALJA
ID   PEPC_CALJA   STANDARD;   PRT;   388 AA.
AC   Q9N2D3;
DT   16-OCT-2001 (Rel. 40, Created)
DT   16-OCT-2001 (Rel. 40, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
GN   PGC.
OS   Callithrix jacchus (Common marmoset).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Platyrrhini; Callitrichidae;
OC   Callithrix.
OX   NCBI_TaxID=9483;
RN   [1]
RP   SEQUENCE FROM N.A., SEQUENCE OF 17-26, FUNCTION, AND ENZYME
RP   REGULATION.
RC   TISSUE=Gastric mucosa;
RX   MEDLINE=20250834; PubMed=10788784;
RA   Kageyama T.;
RT   "New World monkey pepsinogens A and C, and prochymosins. Purification,
RT   characterization of enzymatic properties, cDNA cloning, and molecular
RT   evolution.";
RL   J. Biochem. 127:761-770(2000).
CC   -!- FUNCTION: Hydrolyzes a variety of proteins.
CC   -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC   shows preferential cleavage at Tyr-|-Xaa bonds; high activity
CC   towards hemoglobin as substrate.
CC   -!- ENZYME REGULATION: Inhibited by pepstatin.
CC   -!- MISCELLANEOUS: The optimal pH is around 2.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---
DR   EMBL; AB038385; BAA90872.1; -.
DR   PIR; JC7246; JC7246.
DR   HSSP; P20142; 1AVF.
DR   MEROPS; A01.003; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
FT   SIGNAL   1   16
FT   PROPEP   17   59   ACTIVATION PEPTIDE (BY SIMILARITY).
FT   CHAIN   60   388   GASTRICSIN.
FT   ACT_SITE   91   91   BY SIMILARITY.
FT   ACT_SITE   276   276   BY SIMILARITY.
FT   DISULFID   104   109   BY SIMILARITY.
FT   DISULFID   267   271   BY SIMILARITY.
FT   DISULFID   310   343   BY SIMILARITY.
SQ   SEQUENCE   388 AA; 42503 MW; 0BC48DBD1F7D2D8C CRC64;

Alignment Scores:

Pred. No.:	5.62e-16	Length:	388
Score:	351.50	Matches:	112
Percent Similarity:	45.16%	Conservative:	56
Best Local Similarity:	30.11%	Mismatches:	115
Query Match:	10.53%	Indels:	89
DB:	1	Gaps:	17

US-09-806-194-1 (1-1804) x PEPC_CALJA (1-388)

```
Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
Db         73 TyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThr 92

Qy        334 GGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCGCAC 369
Db         93 GlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHis 112

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        113 Ser------------ArgPheAsnProSerAlaSerSerThrTyrSerSerAsnGlyGln 128

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        129 ThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 148

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
Db        149 ThrVal---------------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSer 163

Qy        550 GAGAATTTCTTTTTGCCTGGG--------------ATTAAATGGAATGGAATACTTGGC 594
Db        164 GluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 181

Qy        595 CTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTG 654
Db        182 LeuAlaTyrProAlaLeuSerMetGlyGlyAlaThr------ThrAlaMetGlnGlyMet 199

Qy        655 GTGACACAAGCAAACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTG 711
Db        200 LeuGlnGluGlyAlaLeuThrSerProValPheSerPheTyrLeuSerAsnGln------ 217

Qy        712 CCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGT 771
Db        218 -----------GlnGlySerSerGlyGlyAlaValIlePheGlyGlyValAspSerSer 233

Qy        772 TTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAA 831
Db        234 LeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGly 253
```

```
Qy        832 ATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCA 891
Db        254 IleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------Gly 270

Qy        892 GACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGAT 951
Db        271 CysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSer 290

Qy        952 GCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGG 1011
Db        291 AlaPheLeuGluAla---------------------------------------------- 295

Qy       1012 ACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC------ 1065
Db        296 ThrGlyAlaGln--------------------GluAspGluTyrGlyGlnPheLeuVal 308

Qy       1066 ----------------------CCTAAAATCTCCATCTACCTGAGAGATGAGAACTCC 1101
Db        309 AsnCysAspSerIleGlnAsnLeuProThrLeuThrPheIleIle--------------- 323

Qy       1102 AGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCC 1161
Db        324 ---AsnGlyValGluPheProLeuProProSerSerTyrIle----------------- 336

Qy       1162 GGCCTGAATTATGAATGTTAC---CGATTCGGCATTTCCCCA----------------- 1200
Db        337 ---LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSerGln 355

Qy       1201 ---TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db        356 AsnSerGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValPhe 375

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
Db        376 AspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 8
PEPC_CAVPO
ID   PEPC_CAVPO     STANDARD;      PRT;   394 AA.
AC   Q64411;
DT   15-JUL-1998 (Rel. 36, Created)
DT   15-JUL-1998 (Rel. 36, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
GN   PGC.
OS   Cavia porcellus (Guinea pig).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Hystricognathi; Caviidae; Cavia.
OX   NCBI_TaxID=10141;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=92355614; PubMed=1644829;
RA   Kageyama T., Ichinose M., Tsukada S., Miki K., Kurokawa K., Koiwai O.,
RA   Tanji M., Yakabe E., Athauda S.B., Takahashi K.;
RT   "Gastric procathepsin E and progastricsin from guinea pig.
RT   Purification, molecular cloning of cDNAs, and characterization of
RT   enzymatic properties, with special reference to procathepsin E.";
RL   J. Biol. Chem. 267:16450-16459(1992).
CC   -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC       shows preferential cleavage at Tyr-|-Xaa bonds; high activity
CC       towards hemoglobin as substrate.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M88652; AAA37053.1; -.
DR   PIR; B43356; B43356.
DR   HSSP; P20142; 1AVF.
DR   MEROPS; A01.003; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AsppproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
FT   SIGNAL       1    16       POTENTIAL.
FT   PROPEP      17    65       ACTIVATION PEPTIDE.
FT   CHAIN       66   394       GASTRICSIN.
FT   ACT_SITE    97    97       BY SIMILARITY.
FT   ACT_SITE   283   283       BY SIMILARITY.
FT   DISULFID   110   115       BY SIMILARITY.
FT   DISULFID   273   277       BY SIMILARITY.
FT   DISULFID   316   349       BY SIMILARITY.
SQ   SEQUENCE   394 AA;  42995 MW;  114F08E105D49865 CRC64;

Alignment Scores:
Pred. No.:                3.54e-14        Length:          394
Score:                    324.50          Matches:         107
Percent Similarity:       46.07%          Conservative:    63
Best Local Similarity:    29.00%          Mismatches:      116
Query Match:              9.72%           Indels:          83
DB:                       1               Gaps:            18

US-09-806-194-1 (1-1804) x PEPC_CAVPO (1-394)

Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
Db         79 TyrPheGlyGlnIleSerLeuGlyThrProProGlnSerPheGlnValLeuPheAspThr 98

Qy        334 GGAAGCAGTAACTTT----------------------GCCGTGGCAGGAACCCCGCAC 369
Db         99 GlySerSerAsnLeuTrpValProSerValTyrCysSerSerLeuAlaCysThrThrHis 118

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        119 ------------ThrArgPheAsnProArgAspSerSerThrTyrValAlaThrAspGln 134

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        135 SerPheSerLeuGluTyrGlyThrGlySerLeuThrGlyValPheGlyTyrAspThrMet 154

Qy        490 ACCATC--------------CCCAAA---GGCTTCAATACTTCTTTTCTTGTCAACATT 531
Db        155 ThrIleGlnAspIleGlnValProLysGlnGluPheGlyLeuSer--------------- 169

Qy        532 GCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGG--------------ATTAAA 576
Db        170 ------------GluThrGlu-----------ProGlySerAspPheValTyrAlaGlu 181

Qy        577 TGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAG 636
Db        182 PheAspGlyIleLeuGlyLeuGlyTyrProGlyLeuSerGluGlyGlyAlaThr------ 199

Qy        637 ACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAG 693
Db        200 ThrAlaMetGlnGlyLeuLeuArgGluGlyAlaLeuSerGlnSerLeuPheSerValTyr 219

Qy        694 ATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCT------GGGACCAACGGAGGTAGTCTT 747
Db        220 Leu-----------------------GlySerGlnGlnGlySerAspGluGlyGlnLeu 231

Qy        748 GTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAG 807
Db        232 IleLeuGlyGlyValAspGluSerLeuTyrThrGlyAspIleTyrTrpThrProValThr 251

Qy        808 GAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAAT 867
Db        252 GlnGluLeuTyrTrpGlnIleGlyIleGluGlyPheLeuIleAspGlySerAlaSerGly 271
```

```
Qy        868 CTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTG 927
                  |||               :::   |||||||||:::|:|||:::||| .|
Db        272 TrpCysSerArg---------GlyCysGlnGlyIleValAspThrGlyThrSerLeuLeu 288

Qy        928 CGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGGCATCTCTGATT 987
                ::|||               :::|||:::|||:::
Db        289 ThrValProSerAspTyrLeuSerThrLeuValGlnAlaIleGlyAlaGluGlu------ 306

Qy        988 CCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAA 1047
                |||:::   :::   |||               :::::|||   :::      :::
Db        307 AsnGluTyrGlyGluTyrPhe--------------ValSerCysSerSerIleGlnAsp 321

Qy       1048 ACACCTTGGTCTTACTTCCCIAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGG 1107
                |||         |||          :::
Db        322 LeuProThrLeuThrPheValIleSerGlyVal------------------------- 332

Qy       1108 TCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCC---------------ATG 1152
                         :::   |||      ||||||                    |||
Db        333 ------GluPheProLeuSerProSerAlaTyrIleLeuSerGlyGluAsnTyrCysMet 350

Qy       1153 ATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAAT--- 1209
              :::|||            |||                  :::||||||
Db        351 ValGlyLeuGluSerThrTyr-----------------ValSerProGlyGlyGlyGlu 364

Qy       1210 ---GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCC 1266
                  :::::::|||         :::      :::|||   ::::::|||   ||
Db        365 ProValTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuAla 384

Qy       1267 CAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
                 ||||||||||||||   :::
Db        385 AsnAsnArgValGlyPheAlaThrAla 393

RESULT 9
RENI_MOUSE
ID   RENI_MOUSE     STANDARD;      PRT;   402 AA.
AC   P06281; P97911; Q62153; Q62154;
DT   01-JAN-1988 (Rel. 06, Created)
DT   01-JAN-1988 (Rel. 06, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Renin precursor, renal (EC 3.4.23.15) (Angiotensinogenase).
GN   REN1 OR REN.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c;
RX   MEDLINE=84182525; PubMed=6370686;
RA   Holm I., Ollo R., Panthier J.-J., Rougeon F.;
RT   "Evolution of aspartyl proteases by gene duplication: the mouse renin
RT   gene is organized in two homologous clusters of four exons.";
RL   EMBO J. 3:557-562(1984).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BALB/c; TISSUE=Kidney;
RX   MEDLINE=90067953; PubMed=2685761;
RA   Kim W.S., Murakami K., Nakayama K.;
RT   "Nucleotide sequence of a cDNA coding for mouse Ren1 preprorenin.";
RL   Nucleic Acids Res. 17:9480-9480(1989).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=DBA/2, and C57BL/10;
RX   MEDLINE=90108722; PubMed=2691339;
RA   Burt D.W., Mullins L.J., George H., Smith G., Brooks J., Pioli D.,
RA   Brammar W.J.;
RT   "The nucleotide sequence of a mouse renin-encoding gene, Ren-1d, and
RT   its upstream region.";
RL   Gene 84:91-104(1989).
RN   [4]
RP   SEQUENCE OF 1-30 FROM N.A.
RC   TISSUE=Kidney;
RX   MEDLINE=84298161; PubMed=6089205;
RA   Panthier J.-J., Dreyfus M., Roux D.T.L., Rougeon F.;
RT   "Mouse kidney and submaxillary gland renin genes differ in their 5'
RT   putative regulatory sequences.";
RL   Proc. Natl. Acad. Sci. U.S.A. 81:5489-5493(1984).
RN   [5]
RP   SEQUENCE OF 1-31 FROM N.A.
RX   MEDLINE=85085936; PubMed=6392850;
RA   Field L.J., Philbrick W.M., Howles P.N., Dickinson D.P.,
RA   McGowan R.A., Gross K.W.;
RT   "Expression of tissue-specific Ren-1 and Ren-2 genes of mice:
RT   comparative analysis of 5'-proximal flanking regions.";
RL   Mol. Cell. Biol. 4:2321-2331(1984).
RN   [6]
RP   SEQUENCE OF 22-37 AND 72-80.
RC   STRAIN=C57BL/10ROS X C3H/HEROS; TISSUE=Kidney;
RX   MEDLINE=97182599; PubMed=9030738;
RA   Jones C.A., Petrovic N., Novak E.K., Swank R.T., Sigmund C.D.,
RA   Gross K.W.;
RT   "Biosynthesis of renin in mouse kidney tumor As4.1 cells.";
RL   Eur. J. Biochem. 243:181-190(1997).
CC   -!- FUNCTION: Renin is a highly specific endopeptidase, whose only
CC       known function is to generate angiotensin I from angiotensinogen
CC       in the plasma, initiating a cascade of reactions that produce an
CC       elevation of blood pressure and increased sodium retention by the
CC       kidney.
CC   -!- CATALYTIC ACTIVITY: Cleaves Leu-|- bond in angiotensinogen to
CC       generate angiotensin I.
CC   -!- TISSUE SPECIFICITY: Kidney.
CC   -!- INDUCTION: RENAL RENIN IS SYNTHESIZED BY THE JUXTAGLOMERULAR CELLS
CC       OF THE KIDNEY IN RESPONSE TO DECREASED BLOOD PRESSURE AND SODIUM
CC       CONCENTRATION.
CC   -!- POLYMORPHISM: In inbred mouse strains, there are at least two
CC       alleles which can occur at the Ren1 locus: Ren-1D and REN-1C.
CC       The sequence shown is that of REN-1C.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; X00810; CAA25391.1; -.
DR   EMBL; X00811; CAA25391.1; JOINED.
DR   EMBL; X00812; CAA25391.1; JOINED.
DR   EMBL; X00813; CAA25391.1; JOINED.
DR   EMBL; X00814; CAA25391.1; JOINED.
DR   EMBL; X00815; CAA25391.1; JOINED.
DR   EMBL; X00816; CAA25391.1; JOINED.
DR   EMBL; X00850; CAA25391.1; JOINED.
DR   EMBL; X00851; CAA25391.1; JOINED.
DR   EMBL; X16642; CAA34636.1; -.
DR   EMBL; K02596; AAA40045.1; -.
DR   EMBL; M32352; AAA40043.1; -.
DR   EMBL; K02800; AAA40044.1; -.
DR   EMBL; M34190; AAA40042.1; -.
DR   PIR; A00989; REMSK.
DR   HSSP; P00796; 1SMR.
DR   MEROPS; A01.007; -.
DR   MGD; MGI:97898; Ren1.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Plasma; Glycoprotein; Zymogen;
KW   Signal.
FT   SIGNAL        1     21
FT   PROPEP       22     71       ACTIVATION PEPTIDE.
FT   CHAIN        72    402       RENIN.
FT   ACT_SITE    102    102       BY SIMILARITY.
```

```
FT   ACT_SITE    287    287       BY SIMILARITY.
FT   DISULFID    115    122       BY SIMILARITY.
FT   DISULFID    278    282       BY SIMILARITY.
FT   CARBOHYD     69     69       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    139    139       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD    320    320       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT      58     58       W -> R (IN REN-1D).
FT   VARIANT      68     68       T -> I (IN REN-1D).
FT   VARIANT     160    160       S -> V (IN REN-1D).
FT   VARIANT     315    315       E -> D (IN REN-1D).
FT   VARIANT     352    352       N -> Y (IN REN-1D).
FT   CONFLICT      6     23       MISSING (IN REF. 1).
FT   CONFLICT     24     24       T -> I (IN REF. 1).
FT   CONFLICT    163    163       V -> VSRV (IN REF. 1).
SQ   SEQUENCE   402 AA;  44342 MW;  D42920B555E97A38 CRC64;

Alignment Scores:
Pred. No.:                7.09e-14      Length:          402
Score:                    320.00        Matches:         126
Percent Similarity:       43.54%        Conservative:    66
Best Local Similarity:    28.57%        Mismatches:      181
Query Match:              9.58%         Indels:          68
DB:                       1             Gaps:            21

US-09-806-194-1 (1-1804) x RENI_MOUSE (1-402)

Qy         28 CTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTC 87
              :::|||||||   |||     ||||||                 :::|||     |||
Db          6 MetProLeuTrpAlaLeuLeuLeuLeu--------------TrpSerProCysThrPhe 23

Qy         88 ACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGA--- 144
              :::|||||||   |||   |||                 :::          |||
Db         21 SerLeuProThrArgThrAlaThrPheGluArgIleProLeuLysLysMetProSerVal 40

Qy        145 CCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAG--------- 195
                          ||||||      |||         |||:::   |||
Db         41 ArgGluIleLeuGluGluArgGlyValAspMetThrArgLeuSerAlaGluTrpGlyVal 60

Qy        196 ------------CCTGCC---------CTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTG 234
                         |||:::         |||   ||||||           ||||:: ||
Db         61 PheThrLysArgProSerLeuThrAsnLeuThrSerProValValLeuThrAsnTyrLeu 80

Qy        235 GCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATC 294
                      |||   |||                 ||||||   |||:::   | |
Db         81 ------------AsnThrGln-----------------TyrTyrGlyGluIleGlyIle 90

Qy        295 GGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTG 354
              |||||||||||||||      :::::::::   ||||||||||||||:::|||      |||
Db         91 GlyThrProProGlnThrPheLysValIlePheAspThrGlySerAlaAsnLeuTrpVal 110

Qy        355 GCAGGAACCCCGCACTCCTAC-----------------ATAGACACGTACTTTGACACA 396
                 |||      |||                    |||   :::   :::::::::
Db        111 ProSerThrLysCysSerArgLeuTyrLeuAlaCysGlyIleHisSerLeuTyrGluSer 130

Qy        397 GAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGA 456
                 ||||||:::|||      |||  |||  ||||::   |||      |||
Db        131 SerAspSerSerSerTyrMetGluAsnGlySerAspPheThrIleHisTyrGlySerGly 150

Qy        457 AGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCT 516
                    ||||||:::   :::|||   ||||||:::      |||         :::
Db        151 ArgValLysGlyPheLeuSerGlnAspSerValThrVal---GlyGlyIleThrValThr 169

Qy        517 TTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAA 576
                          |||                  |||   |||         |||
Db        170 GlnThrPheGlyGluValThrGluLeuProLeuIleProPheMetLeu------AlaLys 187

Qy        577 TGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAG 636
              ::::::|||:::|||||||::   :::         |||:::              :::
Db        188 PheAspGlyValLeuGlyMetGlyPhePro------AlaGlnAlaValGlyGlyValThr 205

Qy        637 ACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAG 693
                 ||||||   :::::::::|||      :::         |||||||||:::
Db        206 ProValPheAspHisIleLeuSerGlnGlyValLeuLysGluGluValPheSerValTyr 225

Qy        694 ATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCITG 753
                                        ||||||         ||||||   :::|||||||
Db        226 Tyr------------------AsnArgGlySerHisLeuLeuGlyGlyGluValValLeu 239

Qy        754 GGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAG 813
                 ||||||   :::|||      |||:::|||:::      |||      |||   :::
Db        240 GlyGlySerAspProGlnHisTyrGlnGlyAsnPheHisTyrValSerIleSerLysThr 259

Qy        814 TGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGAC 873
                    :::|||||||   :::      :::   :::|||      |||      |||
Db        260 AspSerTrpGlnIleThrMetLysGlyValSerValGly------SerSerThrLeuLeu 277

Qy        874 TGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCIGCGCCTG 933
                 |||   |||      |||         :::||||||:::|||:::::   :::
Db        278 CysGluGluGlyCysAla------ValValValAspThrGlySerSerPheIleSerAla 295

Qy        934 CCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTG---GCCCGCGCATCTCTGATTCCA 990
                 |||                 :::::::::|||:::   |||:::'        |||
Db        296 ProThrSerSerLeuLysLeuIleMetGlnAlaLeuGlyAlaLysGluLysArgIleGlu 315

Qy        991 GAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACA 1050
                 |||:::                       :::   |||         |||:::
Db        316 GluTyr----------------------ValValAsnCys---------SerGlnVal 324

Qy       1051 CCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCA 1110
                 |||            |||   ||||||      |||                 |||:::
Db        325 Pro------ThrLeuProAspIleSerPheAspLeu-----------GlyGlyArgAla 338

Qy       1111 TTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTG--- 1167
                 :::   ::::::      :::       |||           |||
Db        339 TyrThrLeuSerSerThrAspTyrValLeuGlnTyrProAsnArgArgAspLysLeuCys 358

Qy       1168 AATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTG---GTGATCGGT 1224
                       :::         |||   |||   |||      :::   |||:::|||
Db        359 ThrLeuAlaLeuHisAlaMetAspIleProProProThrGlyProValTrpValLeuGly 378

Qy       1225 GCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTC 1284
                 ||||||   :::      ||||||      ||||||||         |||:::||||||
Db        379 AlaThrPheIleArgLysPheTyrThrGluPheAspArgHisAsnAsnArgIleGlyPhe 398

Qy       1285 GCA 1287
                 |||
Db        399 Ala 399

RESULT 10
CATD_CLUHA
ID   CATD_CLUHA     STANDARD;      PRT;   396 AA.
AC   Q9DEX3;
DT   28-FEB-2003 (Rel. 41, Created)
DT   28-FEB-2003 (Rel. 41, Last sequence update)
DT   15-SEP-2003 (Rel. 42, Last annotation update)
DE   Cathepsin D precursor (EC 3.4.23.5).
OS   Clupea harengus (Atlantic herring).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Actinopterygii; Neopterygii; Teleostei; Clupeomorpha; Clupeidae;
OC   Clupea.
OX   NCBI_TaxID=7950;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Nielsen L.B., Stougaard P., Andersen P.S., Pedersen L.H.;
RT   "Cloning and sequence determination of herring muscle cathepsin D.";
RL   Submitted (OCT-2000) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE OF 62-82.
RC   TISSUE=Skeletal muscle;
RX   MEDLINE=21165469; PubMed=11207447;
RA   Nielsen L.B., Nielsen H.H.;
RT   "Purification and characterization of cathepsin D from herring muscle
RT   (Clupea harengus).";
RL   Comp. Biochem. Physiol. 128B:351-363(2001).
CC   -!- FUNCTION: Cathepsin D is an acid protease active in intracellular
CC       protein breakdown.
```

```
CC   -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC       that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
CC       chain of insulin.
CC   -!- ENZYME REGULATION: Inhibited by pepstatin.
CC   -!- SUBUNIT: Monomer.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- MISCELLANEOUS: The isoelectric point is 6.8. Has optimal activity
CC       at pH 2.5 with hemoglobin as the substrate and the optimal
CC       temperature is 37 degrees Celsius.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; AF312364; AAG27733.1; -.
DR   HSSP; P07339; 1LYB.
DR   MEROPS; A01.009; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen.
FT   SIGNAL       1     18       POTENTIAL.
FT   PROPEP      19     61       ACTIVATION PEPTIDE.
FT   CHAIN       62    396       CATHEPSIN D.
FT   ACT_SITE    94     94       BY SIMILARITY.
FT   ACT_SITE   281    281       BY SIMILARITY.
FT   DISULFID   107    114       BY SIMILARITY.
FT   DISULFID   272    276       BY SIMILARITY.
FT   DISULFID   315    352       BY SIMILARITY.
FT   CARBOHYD   131    131       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   396 AA;  43315 MW;  D0375DC38567A31B CRC64;

Alignment Scores:
Pred. No.:                1.41e-13        Length:          396
Score:                    315.50          Matches:         112
Percent Similarity:       42.86%          Conservative:    65
Best Local Similarity:    27.12%          Mismatches:      141
Query Match:              9.45%           Indels:          95
DB:                       1               Gaps:            18

US-09-806-194-1 (1-1804) x CATD_CLUHA (1-396)

Qy        148 GGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCG 207
              ||||||   :::   :::|||      |||          :::      |||
Db         47 GlyThrAsnSerLeuGlnHisAsnGlnGlyProSerSerAsnAlaPro--------- 83

Qy        208 TCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGC 267
              :::|||          |||::::::
Db         64 ThrProGluThrLeuLysAsnTyrMet----------------------------Asp 78

Qy        268 CGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTT 327
                    ||||||   |||:::   :::||||||||||   |||         ::::::
Db         74 AlaGlnTyrTyrGlyGluIleGlyLeuGlyThrProValGlnMetPheThrValValPhe 93

Qy        328 GACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACG--- 384
              ||||||||||||||||||      :::              ||:::   |||
Db         94 AspThrGlySerSerAsnLeuTrpLeuProSerIleHisCysSerPheThrAspIleAla 113

Qy        385 --------------TACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
                          ::::::      :::|||||||||||||         |||
Db        114 CysLeuLeuHisHisLysTyrAsnGlyAlaLysSerSerThrTyrValLysAsnGlyThr 133

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
              :::      ::::::|||      ||||||   :::|:|::::::   :::|||
Db        134 GluPheAlaIleGlnTyrGlySerGlySerLeuSerGlyTyrLeuSerGlnAspSerCys 153

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
              ||||||          :::|||          :::|||
Db        154 ThrIle-----------------GlyAspIleValValGluLysGlnLeuPhe---Gly 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATT---------------AAATGGAATGGAATACTTGGC 594
              |||          ||||||:::                |||::::::||||||||||||
Db    ▸   167 GluAlaIleLysGlnProGlyValAlaPheIleAlaAlaLysPheAspGlyIleLeuGly 186

Qy        595 CTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTG 654
              :::|||||      ::::::                :::         ||||||   :::
Db        187 MetAlaTyrProArgIleSer------ValAspGlyValProProValPheAspMetMet 204

Qy        655 GTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTG 711
              ::::::::||      :::      ||||||||||||      :::
Db        205 MetSerGlnLysLysValGluGlnAsnValPheSerPheTyrLeu-------------- 219

Qy        712 CCCGTTGCTGGATCTGGGACCAAC---------------GGAGGTAGTCTTGTCTTGGGT 756
                                  |||                ||||||   |||:::||||||
Db        220 ---------------AsnArgAsnProAspThrGluProGlyGlyGluLeuLeuGly 234

Qy        757 GGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGG 816
              |||   :::|||      |||   ||||||      |||   |||:::      :::
Db        235 GlyThrAspProLysTyrTyrThrGlyAspPheAsnTyrValProValThrArgGlnAla 254

Qy        817 TACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGC 876
              |||:::||||||   :::      :::   ||||||   |||   |||   |||   |||
Db        255 TyrTrpGlnIleHisMetAspGlyMetSerIleGlySerGln---LeuThrLeu---Cys 272

Qy        877 AGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCC 936
              ::::::          :::||||||||||||:::||||||:::|||:::      |||
Db        273 LysAsp------GlyCysGluAlaIleValAspThrGlyThrSerLeuIleThrGlyPro 290

Qy        937 CAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATT--------- 987
                    |||:::   :::|||:::             ||||||
Db        291 ProAlaGluValArgAlaLeuGlnLysAlaIleGlyAlaIleProLeuIleGlnGlyGlu 310

Qy        988 --------------------------------CCAGAATTCTCTGATGGTTTCTGGACT 1014
                                               |||      |||      |||
Db        311 TyrMetIleAspCysLysLysValProThrLeuProThrIleSer------PheAsnVal 328

Qy       1015 GGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATC 1074
              |||   :::          |||   |||   |||          |||
Db        329 GlyGlyLys-----------ThrTyrSerLeuThrGlyGluGlnTyr---------- 340

Qy       1075 TCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAG 1134
                   |||:::::          ::::::      ::::::   |||
Db        341 ------ValLeuLysGluSerGlnGlyGlyLysThrIleCysLeuSerGlyLeuMetGly 358

Qy       1135 CTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATT 1194
              |||   |||   |||
Db        359 LeuGluIleProPro---------------------------------------------- 363

Qy       1195 TCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATC 1254
                 |||:::          :::::::|||         :::      :::|||   :::
Db        364 ---ProAlaGlyProLeuTrpIleLeuGlyAspValPheIleGlyGlnTyrTyrThrVal 382

Qy       1255 TTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
              |||||||||          ||||||||||||||   |||
Db        383 PheAspArgGluSerAsnArgValGlyPheAlaLysSer 395

RESULT 11
APR1_ORYSA
ID   APR1_ORYSA     STANDARD;      PRT;   509 AA.
AC   Q42456;
DT   01-NOV-1997 (Rel. 35, Created)
DT   01-NOV-1997 (Rel. 35, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Aspartic proteinase oryzasin 1 precursor (EC 3.4.23.-).
OS   Oryza sativa (Rice).
OC   Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC   Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC   Ehrhartoideae; Oryzeae; Oryza.
OX   NCBI_TaxID=4530;
```

```
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=cv. Nipponbare / Japonica; TISSUE=Seed;
RX   MEDLINE=96048031; PubMed=7556174;
RA   Asakura T., Watanabe H., Abe K., Arai S.;
RT   "Rice aspartic proteinase, oryzasin, expressed during seed ripening
RT   and germination, has a gene organization distinct from those of
RT   animal and microbial aspartic proteinases.";
RL   Eur. J. Biochem. 232:77-83(1995).
CC   -!- DEVELOPMENTAL STAGE: SEED RIPENING AND GERMINATION.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D32165; BAA06876.1; -.
DR   EMBL; D32144; BAA06875.1; -.
DR   PIR; S66516; S66516.
DR   HSSP; P42210; 1QDM.
DR   MEROPS; A01.020; -.
DR   Gramene; Q42456; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   InterPro; IPR000004; SapB.
DR   Pfam; PF00026; asp; 1.
DR   Pfam; PF05184; SapB_1; 1.
DR   Pfam; PF03489; SapB_2; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   ProDom; PD001732; SapB_sub; 1.
DR   SMART; SM00118; SAPB; 2.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Glycoprotein; Signal.
FT   SIGNAL       1    20       POTENTIAL.
FT   PROPEP      21    67       POTENTIAL.
FT   CHAIN       68   509       ASPARTIC PROTEINASE ORYZASIN 1.
FT   DOMAIN     318   416       SPECIFIC TO PLANT ASPARTIC PROTEINASES
FT                              (BY SIMILARITY).
FT   ACT_SITE   103   103       BY SIMILARITY.
FT   ACT_SITE   290   290       BY SIMILARITY.
FT   DISULFID   116   122       BY SIMILARITY.
FT   DISULFID   281   285       BY SIMILARITY.
FT   DISULFID   428   465       BY SIMILARITY.
FT   CARBOHYD   252   252       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   400   400       N-LINKED (GLCNAC. . .) (POTENTIAL).
SQ   SEQUENCE   509 AA;  54145 MW;  182F5DADA4CBE358 CRC64;

Alignment Scores:
Pred. No.:              2.01e-13        Length:         509
Score:                  313.50          Matches:        127
Percent Similarity:     36.59%          Conservative:   75
Best Local Similarity:  23.01%          Mismatches:     179
Query Match:            9.39%           Indels:         171
DB:                     1               Gaps:           19

US-09-806-194-1 (1-1804) x APR1_ORYSA (1-509)

Qy          7 GCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCCG 66
              ::::::|||      ||||||      :::|||   |:|   ||||||   |||:::
Db          5 SerValAlaLeuValLeuLeuAlaAlaValLeuLeuGlnAlaLeuLeuProAlaSerAla 24

Qy         67 GAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTA 126
              |||                  |||      |||              :::.||||
Db         25 GluGluGlyLeuValArgIleAlaLeuLysLysArgProIleAspGluAsnSerArgVal 44

Qy        127 GTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTC 186
                 |||         |||              |||      |||      |||   |::
Db         45 AlaAlaArgLeuSerGly--------------GluGluGlyAlaArgArgLeuGlyLeu 59

Qy        187 GCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGAC 246
                      :::|||           |||      :::   :::|||:::   :::
Db         60 ArgGlyAlaAsnSerLeuGlyGlyGlyGlyGlyGluGlyAspIleValAlaLeuLysAsn 79

Qy        247 AACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCG 306
                 :::                  |||:::   |||:::   ':::||||||||||||
Db         80 TyrMetAsnAlaGln-----------TyrPheGlyGluIleGlyValGlyThrProPro 95

Qy        307 CAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC-------------------- 345
              ||||||      ::::::   ||||||||||||||||||
Db         96 GlnLysPheThrValIlePheAspThrGlySerSerAsnLeuTrpValProSerAlaLys 115

Qy        346 ------TTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAG 399
                  |||:::::|||         ||||||              :::   :::
Db        116 CysTyrPheSerIleAlaCysPhePheHisSer-----------ArgTyrLysSerGly 131

Qy        400 AGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGC 459
              :::|||||||||||||:::      |||              ::::::|||      ||||||
Db        132 GlnSerSerThrTyrGlnLysAsnGlyLysProAlaAlaIleGlnTyrGlyThrGlySer 151

Qy        460 TGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTT 519
                   ||||||      ||||||   |||||||:::
Db        152 IleAlaGlyPhePheSerGluAspSerValThrValGlyAsp----------------- 165

Qy        520 CTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTT-----------TTGCCT 567
                                  :::   :::   :::   |||                |||
Db        166 --------------LeuValValLysAspGlnGluPheIleGluAlaThrLysGluPro 180

Qy        568 GGGATT--------------AAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTT 612
              |||:::              |||::::::|||||||||||||||   :::      :::
Db        181 GlyLeuThrPheMetValAlaLysPheAspGlyIleLeuGlyLeuGlyPheGlnGluIle 200

Qy        613 GCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC 672
              :::      :::                                          :::
Db        201 SerValGlyAspAla-----------------------------------------Val 206

Qy        673 CCCAACGTTTTCTCCATGCAGATGTGTGGA---GCCGGCTTGCCCGTTGCTGGATCTGGG 729
              |||      :::  |.|         |||            ||||||
Db        207 ProValTrpTyrLysMetValGluGlnGlyLeuValSerGluProValPheSerPheTrp 226

Qy        730 ACCAAC--------------------GGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 766
                 |||                    ||||||   :::|||   |||||:::::::|||
Db        227 PheAsnArgHisSerAspGluGlyGluGlyGlyGluIleValPheGlyGlyMetAspPro 246

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
              |||   ||||||||||:::      |||   |:|:::   ::::::   |||:::|||
Db        247 SerHisTyrLysGlyAsnHisThrTyrValProValSerGlnLysGlyTyrTrpGlnPhe 266

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
              |||:::      :::   |||||||||:::::            |||         :::
Db        267 GluMetGlyAspValLeuIleGlyGlyLysThrThrGlyPhe---CysAla------Ser 283

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
                    ||||||   |||||||||||||:::||||||      |||      :::
Db        284 GlyCysSerAlaIleAlaAspSerGlyThrSerLeuLeuAlaGlyProThrAlaIleIle 303

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAA--------------- 993
                  :::   |||   :::                :::::   :::
Db        304 ThrGluIleAsnGluLysIleGlyAlaThrGlyValValSerGlnGluCysLysIhrVal 323

Qy        993 ------------------------------------------------------------ 993

Db        324 ValSerGlnTyrGlyGlnGlnIleLeuAspLeuLeuLeuAlaGluThrGlnProSerLys 343

Qy        993 ------------------------------------------------------------ 993

Db        344 IleCysSerGlnValGlyLeuCysThrPheAspGlyLysHisGlyValSerAlaGlyIle 363

Qy        994 --------------------------TTCTCTGATGGTTTC------------------ 1008
                                        |||:::|||
Db        364 LysSerValValAspAspGluAlaGlyGluSerAsnGlyLeuGlnSerGlyProMetCys 383

Qy       1009 ----------------------TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCG 1044
```

```
                                  |||      :::|||||||||||
Db        384 AsnAlaCysGluMetAlaValValTrpMetGlnAsnGlnLeuAlaGlnAsnLysThrGln 423

Qy       1045 GAAACACCTTGGTCTTAC---------------------------------------- 1062
              :::         :::|||
Db        404 AspLeuIleLeuAsnTyrIleAsnGlnLeuCysAspLysLeuProSerProMetGlyGlu 423

Qy       1063 ----------------------------TTCCCTAAAATCTCCATCTACCTGAGAGAT 1092
                                         |||:::|||||      :::
Db        424 SerSerValAspCysGlySerLeuAlaSerMetProGluIleSerPheThrIleGlyAla 443

Qy       1093 GAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATG 1152
              :::                  :::      :::   ||||::   ||||||
Db        444 Lys------------------LysPheAlaLeuLysProGluGluTyrIleLeu-- Lys 456

Qy       1153 ATGGGGGCCGGCCTGAATTATGAATGTTAC------------CGATTCGGCATTTCCCCA 1200
              :::|||   |||         :::|||                        |||   |||
Db        457 ValGlyGluGlyAlaAlaAlaGlnCysIleSerGlyPheThrAlaMetAspIleProPro 476

Qy       1201 TCCACAAATGCGCTG---GTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
                  |||   :::::::|||       |||      ::::::   :::||
Db        477 ProArgGlyProLeuTrpIleLeuGlyAspValPheMetGlyAlaTyrHisThrValPhe 496

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
              |||      :::   ||||||||||||||||   |||
Db        497 AspTyrGlyLysMetArgValGlyPheAlaLysSer 508

RESULT 12
PEPC_RAT
ID   PEPC_RAT       STANDARD;      PRT;   392 AA.
AC   P04073;
DT   01-NOV-1986 (Rel. 03, Created)
DT   01-NOV-1986 (Rel. 03, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Gastricsin precursor (EC 3.4.23.3) (Pepsinogen C).
GN   PGC.
OS   Rattus norvegicus (Rat).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX   NCBI_TaxID=10116;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=89255508; PubMed=2722863;
RA   Ishihara T., Ichihara Y., Hayano T., Katsura I., Sogawa K.,
RA   Fujii-Kuriyama Y., Takahashi K.;
RT   "Primary structure and transcriptional regulation of rat pepsinogen C
RT   gene.";
RL   J. Biol. Chem. 264:10193-10199(1989).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=Wistar;
RX   MEDLINE=87054020; PubMed=3780741;
RA   Ichihara Y., Sogawa K., Morohashi K., Fujii-Kuriyama Y., Takahashi K.;
RT   "Nucleotide sequence of a nearly full-length cDNA coding for
RT   pepsinogen of rat gastric mucosa.";
RL   Eur. J. Biochem. 161:7-12(1986).
RN   [3]
RP   SEQUENCE OF 16-112.
RC   STRAIN=Wistar;
RX   MEDLINE=84257697; PubMed=6743670;
RA   Arai K.M., Muto N., Tani S., Akahane K.;
RT   "The N-terminal sequence of rat pepsinogen.";
RL   Biochim. Biophys. Acta 788:256-261(1984).
CC   -!- CATALYTIC ACTIVITY: More restricted specificity than pepsin A, but
CC       shows preferential cleavage at Tyr-|-Xaa bonds; high activity
CC       towards hemoglobin as substrate.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M25993; AAA41827.1; -.
DR   EMBL; M25985; AAA41827.1; JOINED.
DR   EMBL; M25986; AAA41827.1; JOINED.
DR   EMBL; M25987; AAA41827.1; JOINED.
DR   EMBL; M25988; AAA41827.1; JOINED.
DR   EMBL; M25989; AAA41827.1; JOINED.
DR   EMBL; M25990; AAA41827.1; JOINED.
DR   EMBL; M25991; AAA41827.1; JOINED.
DR   EMBL; M25992; AAA41827.1; JOINED.
DR   EMBL; X04644; CAA28305.1; -.
DR   PIR; A33510; A24608.
DR   HSSP; P20142; 1AVF.
DR   MEROPS; A01.003; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Zymogen; Digestion; Signal.
FT   SIGNAL       1     16
FT   PROPEP      17     62       ACTIVATION PEPTIDE.
FT   CHAIN       63    392       GASTRICSIN.
FT   ACT_SITE    94     94
FT   ACT_SITE   280    280
FT   DISULFID   107    112       BY SIMILARITY.
FT   DISULFID   270    275       BY SIMILARITY.
FT   DISULFID   314    347       BY SIMILARITY.
FT   CONFLICT    31     31       E -> Q (IN REF. 3).
FT   CONFLICT   103    103       S -> A (IN REF. 3).
FT   CONFLICT   109    109       S -> L (IN REF. 3).
SQ   SEQUENCE   392 AA;  42833 MW;  092A5EAF2783EDD1 CRC64;

Alignment Scores:
Pred. No.:              2.06e-13        Length:         392
Score:                  313.00          Matches:        105
Percent Similarity:     45.22%          Conservative:   56
Best Local Similarity:  29.49%          Mismatches:     139
Query Match:            9.37%           Indels:         56
DB:                     1               Gaps:           16

US-09-806-194-1 (1-1804) x PEPC_RAT (1-392)

Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
              |||:::   |||:::   ||||||||||||||||||          :::|||   ||||||
Db         76 TyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThr 95

Qy        334 GGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCGCAC 369
              |||||||||||:      |||                     |||   |||   |||
Db         96 GlySerSerAsnLeuTrpValSerSerValTyrCysGlnSerGluAlaCysThrThrHis 115

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
              :::               |||:::      :::|||||||||||||   :::::|||
Db        116 Ala------------ArgPheAsnProSerLysSerSerThrTyrTyrThrGluGlyGln 131

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
                  :::::::::|||      ||||||   ||||||||:   |||   |||   :::
Db        132 ThrPheSerLeuGlnTyrGlyThrGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 151

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
              |||:::                 |||      |||                    |||
Db        152 ThrVal--------------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSer 166

Qy        550 GAGAATTTCTITTTGCCTGGG--------------ATTAAATGGAATGGAATACTTGGC 594
              ||||||         ||||||                    :::::::::||||||:::|||
Db        167 GluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 184

Qy        595 CTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTICTTCGACTCCCTG 654
              |||||||||      |||:::      |||          |||          :::
Db        185 LeuAlaTyrProGlyLeuSer------SerGlyGlyAlaThrThrAlaLeuGlnGlyMet 202
```

```
Qy        655 GTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCC 714
Db        203 LeuGlyGlu-----------GlyAlaLeuSerGlnProLeuPheGlyValTyrLeu--- 217

Qy        715 GTTGCTGGATCT------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db        218 ------GlySerGlnGlnGlySerAsnGlyGlyGlnIleValPheGlyGlyValAspLys 235

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db        236 AsnLeuTyrThrGlyGluIleThrTrpValProValThrGlnGluLeuTyrTrpGlnIle 255

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        256 ThrIleAspAspPheLeuIleGlyAspGlnAlaSerGlyTrp---CysSerSerGlnGly 274

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db        275 Cys---GlnGlyIleValAspThrGlyThrSerLeuLeuValMetProAlaGlnTyrLeu 293

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        294 SerGluLeuLeuGlnThrIleGlyAlaGlnGlu------GlyGluTyrGlyGluTyrPhe 311

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068
Db        312 --------------ValSerCysAspSerValSerSer--------------LeuPro 321

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        322 ThrLeuSerPheValLeu-----------------AsnGlyValGlnPheProLeuSer 335

Qy       1129 CCTCAGCTTTAC---ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGA 1185
Db        336 ProSerSerTyrIleIleGlnGluAspAsnPheCysMetValGlyLeuGluSerIleSer 355

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
Db        356 LeuThrSerGluSerGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyr 375

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
Db        376 TyrAlaIlePheAspMetGlyAsnAsnLysValGlyLeuAlaThrSer 391
```

```
RESULT 13
PEPE_CHICK
ID   PEPE_CHICK     STANDARD;      PRT;   383 AA.
AC   P16476;
DT   01-AUG-1990 (Rel. 15, Created)
DT   01-AUG-1990 (Rel. 15, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Embryonic pepsinogen precursor (EC 3.4.23.-).
OS   Gallus gallus (Chicken).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC   Gallus.
OX   NCBI_TaxID=9031;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=88227903; PubMed=3131317;
RA   Hayashi K., Agata K., Mochii M., Yasuqi S., Eguchi G., Mizuno T.;
RT   "Molecular cloning and the nucleotide sequence of cDNA for embryonic
RT   chicken pepsinogen: phylogenetic relationship with prochymosin.";
RL   J. Biochem. 103:290-296(1988).
CC   -!- DEVELOPMENTAL STAGE: SPECIFICALLY SECRETED DURING THE EMBRYONIC
CC       PERIOD IN THE CHICKEN PROVENTRICULUS (GLANDULAR STOMACH).
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; D00215; BAA00153.1; -.
DR   PIR; A41443; A41443.
DR   HSSP; P00794; 4CMS.
DR   MEROPS; A01.028; -.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Digestion; Signal; Glycoprotein.
FT   SIGNAL       1    16       POTENTIAL.
FT   CHAIN       17   383       EMBRYONIC PEPSINOGEN.
FT   ACT_SITE    94    94       BY SIMILARITY.
FT   ACT_SITE   276   276       BY SIMILARITY.
FT   DISULFID   107   112       BY SIMILARITY.
FT   DISULFID   267   271       BY SIMILARITY.
FT   DISULFID   310   344       BY SIMILARITY.
FT   CARBOHYD   132   132       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   204   204       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   309   309       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   CARBOHYD   350   350       N-LINKED (GLCNAC. . .) (POTENTIAL).
FT   VARIANT     51    51       T -> S.
SQ   SEQUENCE   383 AA;  41719 MW;  1642796871611F54 CRC64;
```

```
Alignment Scores:
Pred. No.:                3.26e-13        Length:          383
Score:                    310.00          Matches:         106
Percent Similarity:       42.78%          Conservative:    63
Best Local Similarity:    26.84%          Mismatches:      136
Query Match:              9.28%           Indels:          90
DB:                       1               Gaps:            15

US-09-806-194-1 (1-1804) x PEPE_CHICK (1-383)

Qy        166 CACGCC------GACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGC 219
Db         55 HisAlaProAspValLeuThrValValThrGluProLeuLeu-------------- 69

Qy        220 GCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC 279
Db         70 ------AsnThrLeuAspMet----------------------------GluTyrTyr 77

Qy        280 CTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGC 339
Db         78 GlyThrIleSerIleGlyThrProProGlnAspPheThrValValPheAspThrGlySer 97

Qy        340 AGTAACTTTGCCGTGGCAGGA----------ACCCCGCACTCCTACATAGACACGTAC 387
Db         98 SerAsnLeuTrpValProSerValSerCysThrSerProAlaCysGlnSerHisGlnMet 117

Qy        388 TTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTAC 447
Db        118 PheAsnProSerGlnSerSerThrTyrLysSerThrGlyGlnAsnLeuSerIleHisTyr 137

Qy        448 ACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTC 507
Db        138 GlyThrGlyAspMetGluGlyThrValGlyCysAspThrValThrValAlaSerLeuMet 157

Qy        508 AATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCT 567
Db        158 AspThrAsnGlnLeuPheGlyLeuSerThr---SerGluProGlyGlnPhePheVal--- 175

Qy        568 GGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGT 627
Db        176 TyrValLysPheAspGlyIleLeuGlyLeuGlyTyrProSerLeuAla------AlaAsp 193

Qy        628 TCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTC 684
Db        194 GlyIleThrProValPheAspAsnMetValAsnGluSerLeuLeuGluGlnAsnLeuPhe 213
```

```
Qy        685 TCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGT 744
              |||:::   :::          :::                                |||
Db        214 SerValTyrLeuSerArgGluProMet----------------------GlySerMet 225

Qy        745 CTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATT 804
              :::|||   |||||||||:::   |||   :::   |||   |||   :::   | ::::
Db        226 ValValPheGlyGlyIleAspGluSerTyrPheThrGlySerIleAsnTrpIleProVal 245

Qy        805 AAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTT 864
                  :::   |||:::||||||   :::      :::   :::      |||   :::
Db        246 SerTyrGlnGlyTyrTrpGlnIleSerMetAspSerIleIleValAsnLysGlnGluIle 265

Qy        865 AATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTG 924
                              :::      :::|||||||:::|||:::||||||'::: :|
Db        266 AlaCysSer------------SerGlyCysGlnAlaIleIleAspThrGlyThrSerLeu 287

Qy        925 CTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTG 984
              :::      |||           :::   :::      ||||'|
Db        282 ValAlaGlyProAlaSerAspIleAsnAspIleGlnSerAlaValGly------------ 297

Qy        985 ATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCG 1044
                                                                  .||
Db        298 ----------------------------------------------------AlaAsnGln 309

Qy       1045 GAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGC 1104
                 |||         |||                                  |||
Db        301 AsnThrTyrGlyGluTyr----------------------------------SerVal 309

Qy       1105 AGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGC 1164
                        |||      :::|:|   :::             :::::::|||   |||
Db        309 AsnCysSerHisIleLeuAlaMetProAspValValPhe------ValIleGly---Gly 325

Qy       1165 CTGAATTAT----------------------------------------GAATGTTAC 1182
              :::   |||                                        ||
Db        326 IleGlnTyrProValProAlaLeuAlaTyrThrGluGlnAsnGlyGlnGlyThrCysMet 345

Qy       1183 CGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGC 1242
                        :::   |||   :::      :::::::|||      :::
Db        346 SerSerPheGlnAsnSerSerAlaAspLeuTrpIleLeuGlyAspValPheIleArgVal 365

Qy       1243 TTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCA 1287
              :::|||   |||||||||||||||      ||||||||||   |||
Db        366 TyrTyrSerIlePheAspArgAlaAsnAsnArgValGlyLeuAla 380
```

RESULT 14
CATD_HUMAN
ID CATD_HUMAN STANDARD; PRT; 412 AA.
AC P07339;
DT 01-APR-1988 (Rel. 07, Created)
DT 01-APR-1988 (Rel. 07, Last sequence update)
DT 15-SEP-2003 (Rel. 42, Last annotation update)
DE Cathepsin D precursor (EC 3.4.23.5).
GN CTSD.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=85270436; PubMed=3927292;
RA Faust P.L., Kornfeld S., Chirgwin J.M.;
RT "Cloning and sequence analysis of cDNA for human cathepsin D.";
RL Proc. Natl. Acad. Sci. U.S.A. 82:4910-4914(1985).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=87231068; PubMed=3588310;
RA Westley B.R., May F.E.B.;
RT "Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive
RT human breast cancer cells.";
RL Nucleic Acids Res. 15:3773-3786(1987).
RN [3]
RP SEQUENCE FROM N.A.
RX MEDLINE=91299158; PubMed=2069717;
RA Redecker B., Heckendorf B., Grosch H.W., Mersmann G., Hasilik A.;
RT "Molecular organization of the human cathepsin D gene.";
RL DNA Cell Biol. 10:423-431(1991).
RN [4]
RP SEQUENCE FROM N.A.
RC TISSUE=Kidney;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Ketteman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length
RT human and mouse cDNA sequences.";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [5]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=94085791; PubMed=8262386;
RA May F.E., Smith D.J., Westley B.R.;
RT "The human cathepsin D-encoding gene is transcribed from an estrogen-
RT regulated and a constitutive start point.";
RL Gene 134:277-282(1993).
RN [6]
RP SEQUENCE OF 1-22 FROM N.A.
RX MEDLINE=95021301; PubMed=7935485;
RA Augereau P., Miralles F., Cavailles V., Gaudelet C., Parker M.,
RA Rochefort H.;
RT "Characterization of the proximal estrogen-responsive element of
RT human cathepsin D gene.";
RL Mol. Endocrinol. 8:693-703(1994).
RN [7]
RP SEQUENCE OF 170-180.
RC TISSUE=Liver;
RA Hochstrasser D.F., Frutiger S., Paquet N., Bairoch A., Ravier F.,
RA Pasquali C., Sanchez J.-C., Tissot J.-D., Bjellqvist B., Vargas R.,
RA Appel R.D., Hughes G.J.;
RL Submitted (JUN-1992) to the SWISS-PROT data bank.
RN [8]
RP VARIANT VAL-58.
RX MEDLINE=20179010; PubMed=10716266;
RA Papassotiropoulos A., Bagli M., Kurz A., Kornhuber J., Forstl H.,
RA Maier W., Pauls J., Lautenschlager N., Heun R.;
RT "A genetic variation of cathepsin D is a major risk factor for
RT Alzheimer's disease.";
RL Ann. Neurol. 47:399-403(2000).
RN [9]
RP X-RAY CRYSTALLOGRAPHY (3 ANGSTROMS).
RC TISSUE=Spleen;
RX MEDLINE=93223670; PubMed=8467789;
RA Metcalf P., Fusek M.;
RT "Two crystal structures for cathepsin D: the lysosomal targeting
RT signal and active site.";
RL EMBO J. 12:1293-1302(1993).
RN [10]
RP X-RAY CRYSTALLOGRAPHY (2.5 ANGSTROMS).
RC TISSUE=Liver;
RX MEDLINE=93342076; PubMed=8393577;
RA Baldwin E.T., Bhat T.N., Gulnik S., Hosur M.V., Sowder R.C. II,
RA Cachau R.E., Collins J., Silva A.M., Erickson J.W.;
RT "Crystal structures of native and inhibited forms of human cathepsin

```
RT   D: implications for lysosomal targeting and drug design.";
RL   Proc. Natl. Acad. Sci. U.S.A. 90:6796-6800(1993).
CC   -!- FUNCTION: Acid protease active in intracellular protein breakdown.
CC       Involved in the pathogenesis of several diseases such as breast
CC       cancer and possibly Alzheimer's disease.
CC   -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC       that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
CC       chain of insulin.
CC   -!- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- POLYMORPHISM: The Val-58 allele is significantly overrepresented
CC       in demented patients (11.8%) compared with nondemented controls
CC       (4.9%). Carriers of the Val-58 allele have a 3.1-fold increased
CC       risk for developing AD than noncarriers.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------
DR   EMBL; M11233; AAB59529.1; -.
DR   EMBL; X05344; CAA28955.1; -.
DR   EMBL; M63138; AAA51922.1; -.
DR   EMBL; M63134; AAA51922.1; JOINED.
DR   EMBL; M63135; AAA51922.1; JOINED.
DR   EMBL; M63136; AAA51922.1; JOINED.
DR   EMBL; M63137; AAA51922.1; JOINED.
DR   EMBL; BC016320; AAH16320.1; -.
DR   EMBL; L12980; AAA16314.1; -.
DR   EMBL; S74689; AAD14156.1; -.
DR   EMBL; S52557; AAD13868.1; -.
DR   PIR; A25771; KHHUD.
DR   PDB; 1LYA; 31-JAN-94.
DR   PDB; 1LYB; 31-JAN-94.
DR   PDB; 1LYW; 22-JUL-99.
DR   MEROPS; A01.009; -.
DR   SWISS-2DPAGE; P07339; HUMAN.
DR   Siena-2DPAGE; P07339; -.
DR   Genew; HGNC:2529; CTSD.
DR   MIM; 116840; -.
DR   GO; GO:0004192; F:cathepsin D activity; TAS.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Signal; Zymogen;
KW   Polymorphism; Alzheimer's disease; 3D-structure.
FT   SIGNAL       1     18
FT   PROPEP      19     64       ACTIVATION PEPTIDE.
FT   CHAIN       65    412       CATHEPSIN D.
FT   CHAIN       65    161       CATHEPSIN D LIGHT CHAIN (PROBABLE).
FT   CHAIN      169    412       CATHEPSIN D HEAVY CHAIN (PROBABLE).
FT   ACT_SITE    97     97
FT   ACT_SITE   295    295
FT   DISULFID    91    160
FT   DISULFID   110    117
FT   DISULFID   286    290
FT   DISULFID   329    366
FT   CARBOHYD   134    134       N-LINKED (GLCNAC. . .).
FT   CARBOHYD   263    263       N-LINKED (GLCNAC. . .).
FT   VARIANT     58     58       A -> V (ASSOCIATED WITH INCREASED RISK IN
FT                               AD; POSSIBLY INFLUENCES SECRETION AND
FT                               INTRACELLULAR MATURATION; dbSNP:17571).
FT                               /FTId=VAR_011621.
FT   STRAND      67     74
FT   TURN        75     77
FT   STRAND      78     85
FT   TURN        86     89
FT   STRAND      90     97
FT   TURN        98     99
FT   STRAND     103    107
FT   TURN       108    109
FT   TURN       112    113
FT   HELIX      115    118
FT   TURN       119    119
FT   STRAND     123    123
FT   HELIX      125    127
FT   TURN       129    130
FT   STRAND     132    141
FT   STRAND     146    158
FT   STRAND     172    184
FT   HELIX      188    192
FT   STRAND     197    200
FT   HELIX      204    206
FT   HELIX      208    210
FT   HELIX      214    220
FT   TURN       221    222
FT   STRAND     228    233
FT   STRAND     243    247
FT   TURN       248    248
FT   HELIX      252    254
FT   STRAND     255    263
FT   STRAND     267    267
FT   TURN       268    269
FT   STRAND     270    279
FT   TURN       280    281

Alignment Scores:
Pred. No.:              4.16e-13      Length:          412
Score:                  308.50        Matches:         121
Percent Similarity:     43.85%        Conservative:    75
Best Local Similarity:  27.07%        Mismatches:      180
Query Match:            9.24%         Indels:          71
DB:                     1             Gaps:            22

US-09-806-194-1 (1-1804) x CATD_HUMAN (1-412)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
              !!!!!!!!!!!              !!!  !!!!!!!!      !!!  !!!
Db          6 LeuLeuProLeuAlaLeu------CysLeuLeuAlaAlaPro------AlaSerAlaLeu 21

Qy         85 TTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGA 144
                 :::!!!!!!            :::      !!!   ::::::         !!:
Db         22 ValArgIleProLeuHisLysPheThrSerIleArgArgThrMetSerGluValGlyGly 41

Qy        145 ------------------CCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTC 186
                               !!!      :::::::::      :::
Db         42 SerValGluAspLeuIleAlaLysGlyProValSerLysTyrSerGlnAlaValProAla 61

Qy        187 GCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGAC 246
                 !!!      :::      !!!          !!:::::
Db         62 ValThrGluGlyProIle------ProGluValLeuLysAsnTyrMet------------ 75

Qy        247 AACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCG 306
                                    !!!!!!   !!:::   !!!!!!!!!!!!!!!
Db         76 ------------------AspAlaGlnTyrTyrGlyGluIleGlyIleGlyThrProPro 89

Qy        307 CAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCG 366
                 !!!         ::::::   !!!!!!!!!!!!!!!!!      !!!
Db         90 GlnCysPheThrValValPheAspThrGlySerSerAsnLeuTrpValProSerIleHis 109

Qy        367 CACTCCTACATAGACACG------------------TACTTTGACACAGAGAGGTCTAGC 408
                    :::!!!                       :::::::::::::::!!!!!!
Db        110 CysLysLeuLeuAspIleAlaCysTrpIleHisHisLysTyrAsnSerAspLysSerSer 129

Qy        409 ACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGC 468
                 !!!!!!         !!!            :::   !!!      !!!!!!   :::!!!
Db        130 ThrTyrValLysAsnGlyThrSerPheAspIleHisTyrGlySerGlySerLeuSerGly 149

Qy        469 TTCGTTGGGGAAGACCTCGTCACCATCCCC------AAAGGCTTCAATACTTCTTTTCTT 522
                 ::::::   :::!!!   !!!::::::!!!              :::   !!!   !!!
```

```
Db        150 TyrLeuSerGlnAspThrValSerValProCysGlnSerAlaSerSerAlaSerAlaLeu 169

Qy        523 ------GTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATT--- 573

Db        170 GlyGlyValLysValGluArgGlnValPheGlyGluAlaThrLysGlnProGlyIleThr 189

Qy        574 -----------AAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCA 621

Db        190 PheIleAlaAlaLysPheAspGlyIleLeuGlyMetAlaTyrProArgIleSer------ 207

Qy        622 TCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAAC 678

Db        208 ValAsnAsnValLeuProValPheAspAsnLeuMetGlnGlnLysLeuValAspGlnAsn 227

Qy        679 GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGA 738

Db        228 IlePheSerPheTyr--------------LeuSerArgAspProAspAlaGlnProGly 242

Qy        739 GGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACC 798

Db        243 GlyGluLeuMetLeuGlyGlyThrAspSerLysTyrTyrLysGlySerLeuSerTyrLeu 262

Qy        799 CCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAA 858

Db        263 AsnValThrArgLysAlaTyrTrpGlnValHisLeuAspGlnValGluVal---AlaSer 281

Qy        859 AGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACC 918

Db        282 GlyLeuThrLeu---CysLysGlu------GlyCysGluAlaIleValAspThrGlyThr 298

Qy        919 ACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCA 978

Db        299 SerLeuMetValGlyPro---------ValAspGluValArgGluLeuGlnLysAlaIle 315

Qy        979 TCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACG 1038

Db        316 GlyAlaValProLeuIleGlnGlyGluTyr----------MetIleProCys------ 329

Qy       1039 AATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAAC 1098

Db        330 ---------GluLysValSerThrLeuProAlaIleThrLeuLysLeu---------- 342

Qy       1099 TCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGG 1158

Db        343 GlyGlyLysGlyTyrLysLeuSer------ProGluAspTyrThrLeuLysValSerGln 360

Qy       1159 GCCGGCCTGAATTATGAATGTTACCGATTC---------GGCATTTCCCCATCCACAAAT 1209

Db        361 AlaGlyLysThrLeuCysLeuSerGlyPheMetGlyMetAspIleProProProSerGly 380

Qy       1210 GCGCTG---GTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCC 1266

Db        381 ProLeuTrpIleLeuGlyAspValPheIleGlyArgTyrTyrThrValPheAspArgAsp 400

Qy       1267 CAGAAGAGGGTGGGCTTCGCA 1287

Db        401 AsnAsnArgValGlyPheAla 407

RESULT 15
CATD_MOUSE
ID   CATD_MOUSE     STANDARD;      PRT;   410 AA.
AC   P18242;
DT   01-NOV-1990 (Rel. 16, Created)
DT   01-NOV-1990 (Rel. 16, Last sequence update)
DT   28-FEB-2003 (Rel. 41, Last annotation update)
DE   Cathepsin D precursor (EC 3.4.23.5).
GN   CTSD.
OS   Mus musculus (Mouse).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX   NCBI_TaxID=10090;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J; TISSUE=Brain;
RX   MEDLINE=91088345; PubMed=2263503;
RA   Diedrich J.F., Staskus K.A., Retzel E.F., Haase A.T.;
RT   "Nucleotide sequence of a cDNA encoding mouse cathepsin D.";
RL   Nucleic Acids Res. 18:7184-7184(1990).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=90326544; PubMed=2374732;
RA   Grusby M.J., Mitchell S.C., Glimcher L.H.;
RT   "Molecular cloning of mouse cathepsin D.";
RL   Nucleic Acids Res. 18:4008-4008(1990).
RN   [3]
RP   SEQUENCE FROM N.A.
RC   STRAIN=C57BL/6J;
RX   MEDLINE=94280622; PubMed=8011168;
RA   Hetman M., Perschl A., Saftig P., von Figura K., Peters C.;
RT   "Mouse cathepsin D gene: molecular organization, characterization of
RT   the promoter, and chromosomal localization.";
RL   DNA Cell Biol. 13:419-427(1994).
CC   -!- FUNCTION: Acid protease active in intracellular protein breakdown.
CC   -!- CATALYTIC ACTIVITY: Specificity similar to, but narrower than,
CC       that of pepsin A. Does not cleave the 4-Gln-|-His-5 bond in B
CC       chain of insulin.
CC   -!- SUBUNIT: CONSISTS OF A LIGHT CHAIN AND A HEAVY CHAIN.
CC   -!- SUBCELLULAR LOCATION: Lysosomal.
CC   -!- SIMILARITY: BELONGS TO PEPTIDASE FAMILY A1.
CC   ---------------------------------------------------------------------------
CC   This SWISS-PROT entry is copyright. It is produced through a collaboration
CC   between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC   the European Bioinformatics Institute. There are no restrictions on its
CC   use by non-profit institutions as long as its content is in no way
CC   modified and this statement is not removed. Usage by and for commercial
CC   entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC   or send an email to license@isb-sib.ch).
CC   ---------------------------------------------------------------------------
DR   EMBL; X53337; CAA37423.1; -.
DR   EMBL; X52886; CAA37067.1; -.
DR   EMBL; X68378; CAA48453.1; -.
DR   EMBL; X68379; CAA48453.1; JOINED.
DR   EMBL; X68380; CAA48453.1; JOINED.
DR   EMBL; X68381; CAA48453.1; JOINED.
DR   EMBL; X68382; CAA48453.1; JOINED.
DR   EMBL; X68383; CAA48453.1; JOINED.
DR   PIR; I48278; KHMSD.
DR   HSSP; P07339; 1LYB.
DR   MEROPS; A01.009; -.
DR   MGD; MGI:88562; Ctsd.
DR   InterPro; IPR001969; Aspprotease_site.
DR   InterPro; IPR001461; AspproteaseA1.
DR   Pfam; PF00026; asp; 1.
DR   PRINTS; PR00792; PEPSIN.
DR   PROSITE; PS00141; ASP_PROTEASE; 2.
KW   Hydrolase; Aspartyl protease; Glycoprotein; Lysosome; Zymogen; Signal.
FT   SIGNAL       1     20       POTENTIAL.
FT   PROPEP      21     64       ACTIVATION PEPTIDE (POTENTIAL).
FT   CHAIN       65    410       CATHEPSIN D.
FT   ACT_SITE    97     97       BY SIMILARITY.
FT   ACT_SITE   293    293       BY SIMILARITY.
FT   DISULFID    91    160       BY SIMILARITY.
FT   DISULFID   110    117       BY SIMILARITY.
FT   DISULFID   284    288       BY SIMILARITY.
FT   DISULFID   327    364       BY SIMILARITY.
FT   CARBOHYD   134    134       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
FT   CARBOHYD   261    261       N-LINKED (GLCNAC. . .) (BY SIMILARITY).
SQ   SEQUENCE   410 AA;  44954 MW;  DC4928EC46928BF0 CRC64;

Alignment Scores:
Pred. No.:                5.64e-13        Length:          410
Score:                    306.50          Matches:         103
Percent Similarity:       44.53%          Conservative:    64
Best Local Similarity:    27.47%          Mismatches:      123
Query Match:              9.18%           Indels:          85
DB:                       1               Gaps:            15
```

US-09-806-194-1 (1-1804) x CATD_MOUSE (1-410)

```
Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
Db         79 TyrTyrGlyAspIleGlyIleGlyThrProProGlnCysPheThrValValPheAspThr 98

Qy        334 GGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACG--------- 384
Db         99 GlySerSerAsnLeuTrpValProSerIleHisCysLysIleLeuAspIleAlaCysTrp 118

Qy        385 ---------TACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTC 435
Db        119 ValHisHisLysTyrAsnSerAspLysSerSerThrTyrValLysAsnGlyThrSerPhe 138

Qy        436 ACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATC 495
Db        139 AspIleHisTyrGlySerGlySerLeuSerGlyTyrLeuSerGlnAspThrValSerVal 158

Qy        496 CCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT---------ATTTTTGAA 546
Db        159 ProCysLysSerAspGlnSerLysAlaArgGlyIleLysValGluLysGlnIlePhe--- 177

Qy        547 TCAGAGAATTTCTTTTTGCCTGGGATT--------------AAATGGAATGGAATACTT 591
Db        178 GlyGluAlaThrLysGlnProGlyIleValPheValAlaAlaLysPheAspGlyIleLeu 197

Qy        592 GGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC 651
Db        198 GlyMetGlyTyrProHisIleSer------ValAsnAsnValLeuProValPheAspAsn 215

Qy        652 CTGGTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGC 708
Db        216 LeuMetGlnGlnLysLeuValAspLysAsnIlePheSerPheTyr--------------- 230

Qy        709 TTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db        231 LeuAsnArgAspProGluGlyGlnProGlyGlyGluLeuMetLeuGlyGlyThrAspSer 250

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db        251 LysTyrTyrHisGlyGluLeuSerTyrLeuAsnValThrArgLysAlaTyrTrpGlnVal 270

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        271 HisMetAspGlnLeuGluValGlyAsnGlu---LeuThrLeu---CysLys------Gly 286

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db        287 GlyCysGluAlaIleValAspThrGlyThrSerLeuLeuValGlyProValGluGluVal 306

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        307 LysGluLeuGlnLysAlaIleGlyAlaValProLeuIle---------- ----- 319

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTGCCT 1068
Db        319 ------------------------------------------------------------ 319

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        320 -----------------GlnGlyGluTyrMetIleProCysGluLysValSerSerLeu 338

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGT--------- 1179
Db        334 ProThrValTyrLeuLys------LeuGly---GlyLysAsnTyrGluLeuHisProAsp 350

Qy       1180 ---TACCGATTCGGCATTTCC---------------------------------------- 1197
Db        351 LysTyrIleLeuLysValSerGlnGlyGlyLysThrIleCysLeuSerGlyPheMetGly 370

Qy       1198 --------------CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGC 1242
Db        371 MetAspIleProProProSerGlyProLeuTrpIleLeuGlyAspValPheIleGlySer 390

Qy       1243 TTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCA 1287
Db        391 TyrTyrThrValPheAspArgAspAsnAsnArgValGlyPheAla 405
```

Search completed: September 25, 2003, 11:01:09
Job time : 56.5 secs

GenCore version 5.1.6

DM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:52:04 ; Search time 68 Seconds
(without alignments)
5102.597 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggcccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 283308 seqs, 96168682 residues

Total number of hits satisfying chosen parameters: 566616

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095455_11030/app_query.fasta_1.1991
-DB=PIR_76 -QFMT=fastan -SUFFIX=rpr -MINMATCH=0.1 -LOOPCL=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi -LIST=45
-DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09806194_@CGN_1_1_95_@runat_25092003_095455_11030 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : PIR_76:*
1: pir1:*
2: pir2:*
3: pir3:*
4: pir4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	1178.5	35.3	501	2	A59090	aspartic proteinas
2	367.5	11.0	383	2	JC7573	pepsinogen C - Afr
3	363.5	10.9	377	1	PEMQCJ	gastricsin (EC 3.4
4	355.5	10.6	384	2	A39314	gastricsin (EC 3.4
5	355	10.6	389	2	JE0371	pepsin C (EC 3.4.2
6	353	10.6	388	2	A29937	gastricsin (EC 3.4
7	351.5	10.5	388	2	JC7246	pepsinogen C - com
8	324.5	9.7	394	2	B43356	gastricsin (EC 3.4
9	320	9.6	385	2	JC7575	pepsinogen A - bul
10	320	9.6	402	1	REMSK	renin (EC 3.4.23.1
11	313.5	9.4	509	2	S66516	oryzasin (EC 3.4.2
12	313	9.4	392	1	A24608	gastricsin (EC 3.4
13	310	9.3	383	2	A41443	pepsin (EC 3.4.23.
14	308.5	9.2	412	1	KHHUD	cathepsin D (EC 3.
15	306.5	9.2	410	1	KHMSD	cathepsin D (EC 3.
16	305.5	9.1	401	1	REMSS	renin (EC 3.4.23.1
17	305	9.1	384	2	JC7574	pepsinogen A - Afr
18	305	9.1	407	1	KHRTD	cathepsin D (EC 3.
19	302	9.0	405	2	A25379	saccharopepsin (EC
20	301.5	9.0	398	2	S66465	cathepsin E (EC 3.
21	300.5	9.0	387	2	C38302	pepsin (EC 3.4.23.
22	299	9.0	398	2	I51185	cathepsin D (EC 3.
23	298.5	8.9	387	2	D38302	pepsin (EC 3.4.23.
24	298.5	8.9	391	2	A43356	cathepsin E (EC 3.
25	298.5	8.9	400	2	I47099	renin (EC 3.4.23.1
26	297	8.9	388	1	PEHU	pepsin A (EC 3.4.2
27	296	8.9	388	2	A30142	pepsin A (EC 3.4.2
28	296	8.9	388	2	B30142	pepsin A (EC 3.4.2
29	294.5	8.8	388	1	S19684	pepsin A (EC 3.4.2
30	292	8.7	506	2	T07915	probable aspartic
31	291.5	8.7	632	2	T45858	hypothetical prote
32	291	8.7	388	1	S19682	pepsin A (EC 3.4.2
33	291	8.7	402	1	RERTK	renin (EC 3.4.23.1
34	291	8.7	406	1	REHUK	renin (EC 3.4.23.1
35	290.5	8.7	396	2	S36865	cathepsin E (EC 3.
36	289	8.7	387	2	E38302	pepsin (EC 3.4.23.
37	288	8.6	387	2	B38302	pepsin (EC 3.4.23.
38	288	8.6	388	1	PEMQAJ	pepsin A (EC 3.4.2
39	287	8.6	396	2	A34401	cathepsin E (EC 3.
40	286.5	8.6	334	2	JC4870	pepsin A (EC 3.4.2
41	286	8.6	382	1	PECH	pepsin A (EC 3.4.2
42	286	8.6	388	1	PEMQAR	pepsin A (EC 3.4.2
43	285.5	8.6	387	2	JC7245	pepsinogen A - com
44	285	8.5	396	2	T47207	aspartic proteinas
45	284.5	8.5	386	1	PEPG	pepsin A (EC 3.4.2

ALIGNMENTS

RESULT 1
A59090
aspartic proteinase (EC 3.4.23.-) BACE precursor - human
N;Alternate names: beta-secretase; beta-site APP cleaving enzyme
C;Species: Homo sapiens (man)
C;Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 11-May-2000
C;Accession: A59090
R;Vassar, R.; Bennett, B.D.; Babu-Khan, S.; Kahn, S.; Mendiaz, E.A.; Denis, P.; Teplo
M.A.; Biere, A.L.; Curran, E.; Burgess, T.; Louis, J.C.; Collins, F.; Treanor, J.; Ro
Science 286, 735-741, 1999
A;Title: beta-Secretase cleavage of Alzheimer's amyloid precursor protein by the tran
A;Reference number: A59090; MUID:20002972; PMID:10531052
A;Note: submitted to GenBank, September 1999
A;Accession: A59090
A;Status: not compared with conceptual translation
A;Molecule type: mRNA
A;Residues: 1-501 <VAS>
A;Cross-references: GB:AF190725; NID:g6118538; PIDN:AAF04142.1; PID:g6118539
C;Genetics:
A;Gene: BACE
C;Superfamily: beta-secretase
C;Keywords: Alzheimer's disease; aspartic proteinase; brain; glycoprotein; hydrolase;
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-45/Domain: propeptide #status predicted <PRO>
F;46-501/Product: acid proteinase BACE #status predicted <MAT>
F;461-477/Domain: transmembrane #status predicted <TRN>
F;93,289/Active site: Asp #status predicted
F;153,172,223,354/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;330-380/Disulfide bonds: #status predicted

Alignment Scores:

Pred. No.:	1.28e-76	Length:	501
Score:	1178.50	Matches:	240
Percent Similarity:	62.04%	Conservative:	82
Best Local Similarity:	46.24%	Mismatches:	164
Query Match:	35.29%	Indels:	33
DB:	2	Gaps:	9

```
S-09-806-194-1 (1-1804) x A59090 (1-501)

Qy        19 GCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC  78
             |||      ||||||   |||      |||:::                  :::   |||
Db         2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet---------GlyAlaGlyValLeuPro  18

Qy        79 GCGCCCTTCACG---------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC  125
             |||      |||              ||||||||||||
Db        19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly---------  35

Qy       124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG  183
                          |||  |||                 ||||||
Db        36 ----------------------GlyAlaProLeu--------------GlyLeuArg  42

Qy       184 CTCGCCCTGGAG------CCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCC  237
             |||      |||                  |||              ::::|:|:::
Db        43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu  62

Qy       238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG  297
             ||||||||||||||||:::|||  |||||||::|||||||||:::||||||  :::|||
Db        63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly  82

Qy       298 ACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA  357
             :::|||||||||  |||  |||||||||||||||||||||||||||||||||||||||
Db        83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly  102

Qy       358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC  417
                ||||||  ::::::     |||:::      :::  |||||||||||||||
Db       103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg  122

Qy       418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGG  477
                    |||  ||   |||||||||  |||  |||   :::|||,
Db       123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly  142

Qy       478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT  537
             |||||||||:::|||||  |||  ||;  :::        ||||||||
Db       143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla  162

Qy       538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTA  597
             |||  ||||||:::  ||||||:::  |||      |||  ||||||||||||,||
Db       163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu  182

Qy       598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG  657
             |||||||||  :::|||:::|||      |||||||||  ||||||||||||||:|||,
Db       183 AlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal  202

Qy       658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT  717
             |||  :::::::||||||::|||||:::|||:::|||||||||||  |||:::
Db       203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu  222

Qy       718 GCTGGATCT---------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA  768
                 |||            :::  |||||||||:::::::::|||||||||:::
Db       223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis  242

Qy       769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA  828
             ||||:||||  |||  :::|||||||||||||||||:::  ||||||||||.:::::::
Db       243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGlnVal  262

Qy       829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC  888
             |||:::::::::|||||  ||||||  |||  :::|||||:::||||||||
Db       263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn  282

Qy       889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT  948
             ||||||:::||||||||| ||||||||  ||||||||||||:::::|||||||
Db       283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe  302

Qy       949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC  1008
             :::|||  |||:::::::::      ||||||        ::::||  ||||||||||
Db       303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe  322

Qy      1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT  1068
             |||  |||  ||||||  ||||||          ||||||||:::  ||||||
Db       323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro  342

Qy      1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG  1128
                ||||||:::||||||      |||  :::::::::||||||||||||||||||||
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu  362

Qy      1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA  1185
             ||||||  |||:::::::|||:::              :::      :::|||||:::
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys  382

Qy      1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC  1245
             |||  ||||||  |||:::        |||:::||||||  :::|||||||||||
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe  402

Qy      1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA  1305
             ||||||:::|||||||||||||:::{||||||:::|||||||||  |||  |||
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal  422

Qy      1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC  1365
                           :::  :::  |||||||||  |||  |||:::
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp  442

Qy      1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC  1425
                                    |||      |||      ::::::|||  :::
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet  462

Qy      1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGG  1485
                ::::::|||  |||:::  ::::::  :::  |||::::::      :::|||
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg  481

Qy      1486 TGTCAGCGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG  1536
             |||  |||      |||  :::      :::      :::|||  |||,  |||
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu  500

RESULT 2
JC7573
pepsinogen C - African clawed frog
N;Alternate names: progastricsin
C;Species: Xenopus laevis (African clawed frog)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C;Accession: JC7573; PC7118
R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A;Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinog
A;Reference number: JC7573; MUID:21064922; PMID:11134969
A;Contents: Stomach
A;Accession: JC7573
A;Molecule type: mRNA
A;Residues: 1-383 <IKU>
A;Cross-references: DDBJ:AB045379
A;Accession: PC7118
A;Molecule type: protein
A;Residues: 17-68 <IK2>
C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-lik
C;Genetics:
A;Gene: PgC
C;Superfamily: pepsin
C;Keywords: stomach; zymogen

Alignment Scores:
Pred. No.:              2e-18          Length:         383
Score:                  367.50         Matches:        132
Percent Similarity:     44.20%         Conservative:   70
Best Local Similarity:  28.88%         Mismatches:     154
Query Match:            11.01%         Indels:         101
DB:                     2              Gaps:           25

US-09-806-194-1 (1-1804) x JC7573 (1-383)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC  60
              |||      |||      |||||||::         |||::::::   ::::::|||
Db          1 MetLysPheLeuIleLeuAlaLeuValCysLeuGlnLeuSerGluGlyIleIleArg---  19

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC  120
```

```
Db        20 ------------ValProLeuLysLysPheLysSerMetArgGluValMetArgGluAsn 35

Qy       121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db        36 GlyIleLysAlaProLeuValAspProAlaThr---------LysTyrTyrAsnGlnTyr 52

Qy       181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db        53 AlaThrAlaTyrGluPro--------------------LeuSerAsnTyrMetAspMet 65

Qy       241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db        66 ----------------------------SerTyrTyrGlyGluIleSerIleGlyThr 75

Qy       301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        76 ProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsnLeuTrpValAlaSer 95

Qy       361 ACCCCGCACTCCTACATAGACACG--------------------TACTTTGACACAGAG 399
Db        96 Thr---------TyrCysGlnSerGlnAlaCysThrAsnHisProLeuPheAsnProSer 112

Qy       400 AGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGC 459
Db       113 GlnSerSerThrTyrSerSerAsnGlnGlnGlnPheSerLeuGlnTyrGlyThrGlySer 132

Qy       460 TGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTT 519
Db       133 LeuThrGlyIleLeuGlyTyrAspThrValThrIleGln---------------------- 145

Qy       520 CTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTG------------CCT 567
Db       146 ------AsnValAla------IleSerGlnGlnGluPheGlyLeuSerGluThrGluPro 161

Qy       568 GGG---------------ATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTT 612
Db       162 GlyThrAsnPheValTyrAlaGlnPheAspGlyIleLeuGlyLeuAlaTyrProSerIle 181

Qy       613 GCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC 672
Db       182 AlaValGlyGlyAlaThr------ThrValMetGlnGlyMetMet---GlnGlnAsnLeu 198

Qy       673 CCCAAC------GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCT 726
Db       199 LeuAsnGlnProIlePheGlyPheTyrLeuSerGlyGln----------------Ser 212

Qy       727 GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGAC 786
Db       213 SerGlnAsnGlyGlyGluValAlaPheGlyGlyValAspGlnAsnTyrTyrThrGlyGln 232

Qy       787 ATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAA 846
Db       233 IleTyrTrpThrProValThrSerGluThrTyrTrpGlnIleGlyIleGlnGlyPheSer 252

Qy       847 ATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTG 906
Db       253 IleAsnGlyGlnAlaThrGlyTrp---CysSerGln------GlyCysGlnAlaIleVal 269

Qy       907 GACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCT 966
Db       270 AspThrGlyThrSerLeuLeuThrAlaProGlnSerValPheSerSerLeuIleGlnSer 289

Qy       967 GTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTG 1026
Db       290 IleGly----------------------AlaGlnGlnAspGlnAsnGlyGlnTyrVal 301

Qy      1027 GCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTG 1086
Db       302 ValSerCysSerAsnIleGlnAsn-----------LeuProThrIleSerPheThrIle 317

Qy      1087 AGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATT--- 1143
```

```
Db       318 -----------SerGlyValSerPheProLeuPro------ProSerAlaTyrValLeu 331

Qy      1144 CAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCC------ 1197
Db       332 GlnGlnSerSerGly--------------TyrCys---ThrIleGlyIleMetProThr 345

Qy      1198 ------CCATCCACAAATGCG---------CTGGTGATCGGTGCCACGGTGATGGAGGGC 1242
Db       346 TyrLeuProSerGlnAsnGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgGlu 365

Qy      1243 TTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
Db       366 TyrTyrSerValTyrAspLeuGlyAsnAsnGlnValGlyPheAlaThrAla 382

RESULT 3
PEMQCJ
gastricsin (EC 3.4.23.3) precursor - Japanese macaque (fragment)
N;Alternate names: pepsin C
C;Species: Macaca fuscata (Japanese macaque)
C;Date: 13-Aug-1986 #sequence_revision 19-Oct-1995 #text_change 18-Jun-1999
C;Accession: S19683; A00986; A22402; S16066
R;Kageyama, T.; Tanabe, K.; Koiwai, O.
Eur. J. Biochem. 202, 205-215, 1991
A;Title: Development-dependent expression of isozymogens of monkey pepsinogens and st
A;Reference number: S19681; MUID:92037645; PMID:1935977
A;Accession: S19683
A;Molecule type: mRNA
A;Residues: 1-377 <KAG>
A;Cross-references: EMBL:X59754; NID:g38072; PIDN:CAA42426.1; PID:g38073
R;Kageyama, T.; Takahashi, K.
J. Biol. Chem. 261, 4406-4419, 1986
A;Title: The complete amino acid sequence of monkey progastricsin.
A;Reference number: A00986; MUID:86168133; PMID:3514597
A;Accession: A00986
A;Molecule type: protein
A;Residues: 6-330,'V',332-349,'VY',350-377 <KA2>
R;Kageyama, T.; Takahashi, K.
J. Biochem. 97, 1235-1246, 1985
A;Title: Monkey pepsinogens and pepsins. VII. Analysis of the activation process and
 pepsinogens.
A;Reference number: A22402; MUID:85289106; PMID:3928607
A;Accession: A22402
A;Molecule type: protein
A;Residues: 6-65 <KA3>
C;Comment: This enzyme has more restricted specificity than pepsin A.
C;Comment: The enzyme is activated in a two-step process that gives rise to two end p
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach
F;1-5/Domain: signal sequence (fragment) #status predicted <SIG>
F;6-377/Product: progastricsin #status experimental <ZYM>
F;6-45/Domain: activation peptide #status experimental <APT>
F;46-377/Product: Gly-gastricsin #status experimental <MIN>
F;49-377/Product: Ser-gastricsin #status experimental <MAT>
F;31-32/Cleavage site: Phe-Leu (pepsin) #status experimental
F;45-46/Cleavage site: Phe-Gly (pepsin) #status experimental
F;48-49/Cleavage site: Leu-Ser (pepsin) #status experimental
F;80,265/Active site: Asp #status predicted
F;93-98,256-260,299-332/Disulfide bonds: #status experimental

Alignment Scores:
Pred. No.:                3.87e-18        Length:          377
Score:                    363.50          Matches:         118
Percent Similarity:       44.85%          Conservative:    65
Best Local Similarity:    28.92%          Mismatches:      118
Query Match:              10.89%          Indels:          107
DB:                       1               Gaps:            19

US-09-806-194-1 (1-1804) x PEMQCJ (1-377)

Qy        166 CACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCC 225
Db         44 HisPheGlyAspLeuSerValSerTyrGluPro-------------------------- 54
```

```
Qy       226 AACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAG 285
Db        55 ------MetAlaTyrMetAsp--------------------AlaAlaTyrPheGlyGlu 65

Qy       286 ATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC 345
Db        66 IleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsn 85

Qy       346 TTTGCCGTG----------------------GCAGGAACCCCGCACTCCTACATAGAC 381
Db        86 LeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHisSer--------- 102

Qy       382 ACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTG 441
Db       103 ---ArgPheAsnProSerGluSerSerThrTyrSerThrAsnGlyGlnThrPheSerLeu 121

Qy       442 AAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAA 501
Db       122 GlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeuThrVal------ 139

Qy       502 GGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTT 561
Db       140 ---------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSerGluAsn------ 154

Qy       562 TTGCCTGGG----------------ATTAAATGGAATGGAATACTTGGCCTAGCTTATGCC 606
Db       155 GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGlyLeuAlaTyrPro 174

Qy       607 ACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCA 666
Db       175 ThrLeuSerValAspGlyAlaThr------ThrAlaMetGlnGlyMetValGlnGluGly 192

Qy       667 AACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGA 723
Db       193 AlaLeuThrSerProIlePheSerValTyrLeuSerAspGln----------------- 206

Qy       724 TCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGA 783
Db       207 GlnGlySerSerGlyGlyAlaValValPheGlyGlyValAspSerSerLeuTyrThrGly 226

Qy       784 GACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTG 843
Db       227 GlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGlyIleGluGluPhe 246

Qy       844 GAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATC 903
Db       247 LeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------GlyCysGlnAlaIle 263

Qy       904 GTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAA 963
Db       264 ValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSerAlaLeuLeuGln 283

Qy       964 GCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAG 1023
Db       284 Ala-----------------------------------------ThrGlyAlaGln 288

Qy      1024 CTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC------------------ 1065
Db       289 --------------------GluAspGluTyrGlyGlnPheLeuValAsnCysAsnSer 301

Qy      1066 ------------CCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTC 1113
Db       302 IleGlnAsnLeuProThrLeuThrPheIleIle------------------AsnGlyVal 315

Qy      1114 CGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATCGGGGCCGGCCTGAATTAT 1173
Db       316 GluPheProLeuProProSerSerTyrIle--------------------LeuAsnAsn 328

Qy      1174 GAATGTTAC---CGATTCGGCATTTCCCCA--------------------TCCACAAAT 1209
Db       329 AsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerAlaGlnAsnSerGlnPro 348

Qy      1210 GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAG 1269
```

```
Db       349 LeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuSerAsn 368

Qy      1270 AAGAGGGTGGGCTTCGCAGCGAGC 1293
Db       369 AsnArgValGlyPheAlaThrAla 376
```

RESULT 4
A39314
gastricsin (EC 3.4.23.3) precursor - bullfrog
C;Species: Rana catesbeiana (bullfrog)
C;Date: 19-Jun-1992 #sequence_revision 19-Jun-1992 #text_change 22-Jun-1999
C;Accession: A39314
R;Yakabe, E.; Tanji, M.; Ichinose, M.; Goto, S.; Miki, K.; Kurokawa, K.; Ito, H.; Kag
J. Biol. Chem. 266, 22436-22443, 1991
A;Title: Purification, characterization, and amino acid sequences of pepsinogens and
A;Reference number: A39314; MUID:92042186; PMID:1939266
A;Accession: A39314
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-384 <YAK>
A;Cross-references: GB:M73750; NID:g213687; PIDN:AAA49530.1; PID:g213688
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; hydrolase; protein digestion

Alignment Scores:

Pred. No.:	1.46e-17	Length:	384
Score:	355.50	Matches:	120
Percent Similarity:	42.70%	Conservative:	73
Best Local Similarity:	26.55%	Mismatches:	136
Query Match:	10.65%	Indels:	123
DB:	2	Gaps:	21

US-09-806-194-1 (1-1804) x A39314 (1-384)

```
Qy         67 GAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTA 126
Db         12 GlnLeuSerGluGlyIleIleLysValProLeuLysLysPheLysSerMetArgGluVal 31

Qy        127 GTT-----------------GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCAC 168
Db         32 MetArgAspHisGlyIleLysAlaProValValAspProAlaThr---------LysTyr 48

Qy        169 GCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAAC 228
Db         49 TyrAsnAsnPheAlaThrAlaPheGluPro--------------------LeuAlaAsn 61

Qy        229 TTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATG 288
Db         62 TyrMetAspMet------------------------------SerTyrTyrGlyGluIle 71

Qy        289 CTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTT 348
Db         72 SerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThrGlySerSerAsnLeu 91

Qy        349 GCCGTGGCAGGAACCCCGCACTCCTACATAGACACG--------------------TAC 387
Db         92 TrpVal---------ProSerThrTyrCysGlnSerGlnAlaCysThrAsnHisProGln 108

Qy        388 TTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTAC 447
Db        109 PheAsnProSerGlnSerSerSerTyrSerSerAsnGlnGlnGlnPheSerLeuGlnTyr 128

Qy        448 ACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTC 507
Db        129 GlyThrGlySerLeuThrGlyIleLeuGlyTyrAspThrValGlnIleGln--------- 145

Qy        508 AATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTG--- 564
Db        146 -----------------AsnIleAla------IleSerGlnGlnGluPheGlyLeuSer 157

Qy        565 ---------CCTGGG---------------ATTAAATGGAATGGAATACTTGGCCTAGCT 600
```

```
Db        158 ValThrGluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleLeuGlyLeuAla 177

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              |||   :::::|||:::      :::::      |||           ::::::
Db        178 TyrProSerIleAlaGluGlyGlyAlaThr------ThrValMetGlnGlyMetIle--- 194

Qy        661 CAAGCAAACATCCCCAAC------GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCC 714
              |||   |||:::   |||      :::|||:::      :::   |||
Db        195 GlnGlnAsnLeuIleAsnGlnProLeuPheAlaPheTyrLeuSerGly---------- 210

Qy        715 GTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTG 774
                  :::      |||||||||   :::      |||:||::::   :::
Db        211 ---GlnGlnAsnSerGlnAsnGlyGlyGluValAlaPheGlyGlyValAspGlnAsnTyr 229

Qy        775 TATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATT 834
              |||   |||   |||::::::|||||:::      |||   ||:::|||||   |||
Db        230 TyrSerGlyGlnIleTyrTrpThrProValThrSerGluThrTyrTrpGlnIleGlyIle 249

Qy        835 CTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGAC 894
                  :::   ||||||::      |||   :::
Db        250 GlnGlyPheSerValAsnGlyGlnAlaThrGlyTrp---CysSerGln------GlyCys 266

Qy        895 AAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCG 954
              :::   ||||||||||:|||||||:|||||      |||||   |||||   :::
Db        267 GlnGlyIleValAspThrGlyThrSerLeuLeuThrAlaProGlnSerValPheSerSer 286

Qy        955 GTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACT 1014
              :::::::::::::::
Db        287 LeuMetGlnSerIle---------------------------------------------- 291

Qy       1015 GGGTCCCAG----------------------CTGGCGTGCTGCACGAATTCGGAAACA 1050
              |||:::|||                     :::::::|||
Db        292 GlyAlaGlnGlnAspGlnAsnGlyGlnTyrAlaValSerCysSerAsnIleGlnSerLeu 311

Qy       1051 CCTTGGTCTTACTTCCCT----------------------AAAATCTCCATCTACCTG 1086
              |||         |||                          |||   ||||::
Db        312 ProThrIleSerPheThrIleSerGlyValSerProProLeuProProSerAlaTyrVal 331

Qy       1087 AGAGATGAGAACTCC---------AGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTT 1137
                 :::|||||||         :::                |||   |||:||
Db        332 LeuGlnGlnAsnSerGlyTyrCysThrIleGlyIleMetProThrTyrLeuProSerGln 351

Qy       1138 TACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCC 1197
                 ||||||:::
Db        352 AsnGlyGlnProLeuTrp---------------------------------------- 367

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
                        :::::::|||      :::      :::|||      ::::::
Db        368 -----------------IleLeuGlyAspValPheLeuArgGlnTyrTyrSerValTyr 371

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
              |||           :::||||||||||||||:::
Db        372 AspLeuGlyAsnAsnGlnValGlyPheAlaAlaAla 383

RESULT 5
JE0371
pepsin C (EC 3.4.23.-) precursor - chicken
N;Alternate names: pepsinogen C
C;Species: Gallus gallus (chicken)
C;Date: 23-Jul-1999 #sequence_revision 23-Jul-1999 #text_change 11-May-2000
C;Accession: JE0371
R;Sakamoto, N.; Saiga, H.; Yasugi, S.
Biochem. Biophys. Res. Commun. 250, 420-424, 1998
A;Title: Analysis of temporal expression pattern and cis-regulatory sequences of chicken
A;Reference number: JE0370; MUID:98440813; PMID:9753645
A;Accession: JE0371
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-389 <SAK>
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; hydrolase

Alignment Scores:

Pred. No.:                1.58e-17      Length:          389
Score:                    355.00        Matches:         114
Percent Similarity:       43.32%        Conservative:    58
Best Local Similarity:    28.72%        Mismatches:      121
Query Match:              10.63%        Indels:          104
DB:                       2             Gaps:            16

US-09-806-194-1 (1-1804) x JE0371 (1-389)

Qy        223 GCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTG 282
              :::||||||           :::   |||      :::          ||||||
Db         56 SerAsnPheAlaThrAlaTyrGluProLeuAlaAsnAsnMetAspMetSerTyrTyrGly 75

Qy        283 GAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGT 342
              |||:::   |||||||||||||||||||         :::|||   ||||||||||||||
Db         76 GluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThrGlySerSer 95

Qy        343 AACTTTGCCGTGGCAGGAACC----------------------CCGCACTCCTACATA 378
              |||      |||      |||                          |||:::
Db         96 AsnLeuTrpValProSerThrLeuCysGlnSerGlnAlaCysAlaAsnHisAsn------ 113

Qy        379 GACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACA 438
                  ||||||      ||||||||||:::  ::::::          :::
Db        114 ------GluPheAspProAsnGluSerSerThrPheSerThrGlnAspGluPhePheSer 131

Qy        439 GTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCC 498
              ::::::|||      ||||||  ||||||      |||   |||   |||||||||
Db        132 LeuGlnTyrGlySerGlySerLeuThrGlyIlePheGlyPheAspThrValThrIle--- 150

Qy        499 AAAGGCTTCAAT------------------------------------------ACTTCT 516
              :::|||   :::                                          ||||||
Db        151 GlnGlyIleSerIleThrAsnGlnGluPheGlyLeuSerGluThrGluProGlyThrSer 170

Qy        517 ITTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAA 576
              ||||||   :::
Db        171 PheLeuTyrSer------------------------------------------Pro 175

Qy        577 TGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAG 636
              ::::::|||||||||||||||||:::   ::::::         |||:::
Db        176 PheAspGlyIleLeuGlyLeuAlaProSerIle---------SerAlaGlyGlyAla 192

Qy        637 ACCTTCTTCGACTCCCTGGTGACACAAGCAAACATCCCCAAC------GTTTTCTCCATG 690
              |||               :::   |||   |||:::   :::      |||||||||
Db        193 ThrThrValMetGlnLysMetLeuGlnGluAsnLeuLeuAspPheProValPheSerPhe 212

Qy        691 CAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTC 750
              :::   |||                  |||:::   ||||||   ||||||
Db        213 TyrLeuSerGlyGln-----------------GluGlySerGlnGlyGlyGluLeuVal 226

Qy        751 TTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAA 810
              |||||:::::::|||:::||||||   |||   |||   :::||||:|:::   :::
Db        227 PheGlyGlyValAspProAsnLeuTyrThrGlyGlnIleThrTrpThrProValThrGln 246

Qy        811 GAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTG 870
                  |||:::||||||   |||            :::||||||||||||
Db        247 ThrThrTyrTrpGlnIleGlyIleGluAspPheAlaValGlyGlyGlnSerSerGlyTrp 266

Qy        871 GACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGC 930
                 |||   :::             :::   ||||||||||:::|||||||:::||||||
Db        267 ---CysSerGln------GlyCysGlnGlyIleValAspThrGlyThrSerLeuLeuThr 283

Qy        931 CTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCA 990
              :::|||   :::||||||      :::::::::   :::
Db        284 ValProAsnGlnValPheThrGluLeuMetGlnTyrIleGly--------------Ala 298

Qy        991 GAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACA 1050
              :::      |||          :::|||      :::|||      :::||||   |||
Db        299 GlnAlaAspAsp---------SerGlyGlnTyrValAlaSerCysSerAsnIleGlu--- 314

Qy       1051 CCTTGGTCTTACTTCCCTAAAATC---------------------------------- 1074
                        |||   |||   |||
Db        315 ---------TyrMetProThrIleThrPheValIleSerGlyThrSerProProLeuPro 331
```

```
y       1075 ---TCCATCTACCTGAGAGATGAGAACTCC---------AGCAGGTCATTCCGIATCACA 1122
                 |||   |||:::          ||||||         :::              |||
b        332 ProSerAlaTyrMetLeuGlnSerAsnSerAspTyrCysThrValGlyIleGluSerThr 351

y       1123 ATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTAC 1182
                 ||||||           ||||||:::
)b       352 TyrLeuProSerGlnThrGlyGlnProLeuTrp------------------------- 382

)y      1183 CGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGC 1242
                                        :::::::|||         :::
)b       383 --------------------------------IleLeuGlyAspValPheLeuArgVal 371

)y      1243 TTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
             :::|||   |||:::|||           :::||||:|||||   :::
)b       372 TyrTyrSerIleTyrAspMetGlyAsnAsnGlnValGlyPheAlaThrAla 388

RESULT 6
A29937
gastricsin (EC 3.4.23.3) precursor - human
N;Alternate names: pepsin C; pepsinogen C
C;Species: Homo sapiens (man)
C;Date: 17-Oct-1988 #sequence_revision 17-Oct-1988 #text_change 31-Mar-2000
C;Accession: A29937; A31811; PX0028; I54213; A91125; A23458
R;Hayano, T.; Sogawa, K.; Ichihara, Y.; Fujii-Kuriyama, Y.; Takahashi, K.
J. Biol. Chem. 263, 1382-1385, 1988
A;Title: Primary structure of human pepsinogen C gene.
A;Reference number: A29937; MUID:88087276; PMID:3335549
A;Accession: A29937
A;Molecule type: DNA
A;Residues: 1-388 <HAY>
R;Taggart, R.T.; Cass, L.G.; Mohandas, T.K.; Derby, P.; Barr, P.J.; Pals, G.; Bell, G.I.
J. Biol. Chem. 264, 375-379, 1989
A;Title: Human pepsinogen C (progastricsin). Isolation of cDNA clones, localization to c
A;Reference number: A31811; MUID:89079679; PMID:2909526
A;Accession: A31811
A;Molecule type: mRNA
A;Residues: 1-388 <TAG>
A;Cross-references: GB:J04443; NID:g551175; PIDN:AAA60074.1; PID:g551176
R;Athauda, S.B.P.; Tanji, M.; Kageyama, T.; Takahashi, K.
J. Biochem. 106, 920-927, 1989
A;Title: A comparative study on the NH2-terminal amino acid sequences and some other pro
A;Reference number: PX0023; MUID:90130402; PMID:2515193
A;Accession: PX0028
A;Molecule type: protein
A;Residues: 17-101 <ATH>
R;Pals, G.; Azuma, T.; Mohandas, T.K.; Bell, G.I.; Bacon, J.; Samloff, I.M.; Walz, D.A.;
Genomics 4, 137-148, 1989
A;Title: Human pepsinogen C (progastricsin) polymorphism: evidence for a single locus lo
A;Reference number: I54213; MUID:89290840; PMID:2567697
A;Accession: I54213
A;Status: translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-388 <RES>
A;Cross-references: GB:M23077; NID:g189830; PIDN:AAA60063.1; PID:g387015; GB:J03063
A;Note: parts of this sequence, including the amino end and carboxyl ends of the mature
R;Foltmann, B.; Jensen, A.L.
Eur. J. Biochem. 128, 63-70, 1982
A;Title: Human progastricsin. Analysis of intermediates during activation into gastricsi
A;Reference number: A91125; MUID:83079318; PMID:6816595
A;Accession: A91125
A;Molecule type: protein
A;Residues: 17-39,'ED',42-51,'S',53-64 <FOL>
A;Note: pro-form; 29-Leu was also found
A;Note: activation at pH 2 is proposed to involve conformation change, cleavage after Ph
C;Genetics:
A;Gene: GDB:PGC
A;Cross-references: GDB:119485; OMIM:169740
A;Map position: 6p21.3-6p21.1
A;Introns: 20/2; 70/3; 110/1; 149/3; 216/2; 256/2; 305/3; 338/3
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; hydrolase; protein digestion; stomach; zymogen
F;1-16/Domain: signal sequence #status predicted <SIG>
F;17-59/Domain: propeptide #status experimental <PRO>
F;60-388/Product: gastricsin #status experimental <MAT>

Alignment Scores:
Pred. No.:                2.2e-17        Length:          388
Score:                    353.00         Matches:         120
Percent Similarity:       44.79%         Conservative:    65
Best Local Similarity:    29.06%         Mismatches:      120
Query Match:              10.57%         Indels:          108
DB:                       2              Gaps:            21

US-09-806-194-1 (1-1804) x A29937 (1-388)

Qy       154 CCTGCCGAGCGCCACGCCGACGGC---TTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCC 210
                 ||||||   ::::::       |||   |||::::::      ||||||
Db        50 ProAlaTrpLysTyrArgPheGlyAspLeuSerValThrTyrGluPro------------ 65

Qy       211 CCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGC 270
                                     :::|||   :::|||
Db        66 --------------------MetAlaTyrMetAsp--------------------Ala 71

Qy       271 GGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGAC 330
                ||||::   |||:::   ||||||||||||||||||          :::|||   |||
Db        72 AlaTyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAsp 91

Qy       331 ACTGGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCG 366
                |||||||||||||||      |||                     |||   |||
Db        92 ThrGlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSer 111

Qy       367 CACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGC 426
                .|||||                |||:::         ||||||||||||   :::   |||
Db       112 HisSer------------ArgPheAsnProSerGluSerSerThrTyrSerThrAsnGly 127

Qy       427 TTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTC 486
                    :::::::::|||      ||||||   ||||||||   |||   |||
Db       128 GlnThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThr 147

Qy       487 GTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAA 546
                :::|||:::                |||     |||
Db       148 LeuThrVal--------------GlnSerIleGlnValProAsnGlnGluPheGlyLeu 162

Qy       547 TCAGAGAATTTCTTTTTGCCTGGG--------------ATTAAATGGAATGGAATACTT 591
                |||||||         ||||||                  :::::::::||||||:::
Db       163 SerGluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMet 180

Qy       592 GGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCC 651
                ||||||||||||      |||:::         ::::::      |||
Db       181 GlyLeuAlaTyrProAlaLeuSerValAspGluAlaThr------ThrAlaMetGlnGly 198

Qy       652 CTGGTGACACAAGCAAACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGC 708
                :::|||   :::      :::   :::   ||||||||||:::   :::
Db       199 MetValGlnGluGlyAlaLeuThrSerProValPheSerValTyrLeuSerAsnGln--- 217

Qy       709 TTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
                              |||::::::||||||:::::|||   ||||||:::::
Db       218 ---------------GlnGlySerSerGlyGlyAlaValValPheGlyGlyValAspSer 232

Qy       769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
                |||||||||   |||   |||::::::   |||:::   :::|||   |||:::||||||
Db       233 SerLeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIle 252

Qy       829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
                |||   :::      ||||||||||||:::          |||   |||
Db       253 GlyIleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------ 269

Qy       889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
                 :::|||||||||||||:::|||||:::|||||   :::|||||:::
Db       270 GlyCysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMet 289

Qy       949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
                |||:::::::::|||
Db       290 SerAlaLeuLeuGlnAla---------------------------------------- 295

Qy      1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC--- 1065
```

```
                ||||||:::|||                    |||      :::      |||
Db        296 ---ThrGlyAlaGln--------------------GluAspGluTyrGlyGlnPheLeu 307

Qy       1066 -------------------------CCTAAAATCTCCATCTACCTGAGAGATGAGAAC 1098
                                       |||   ::::::      :::
Db        308 ValAsnCysAsnSerIleGlnAsnLeuProSerLeuThrPheIleIle------------ 323

Qy       1099 TCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGG 1158
                              :::   |||      ||||||
Db        324 ------AsnGlyValGluPheProLeuProProSerSerTyrIle---------------- 336

Qy       1159 GCCGGCCTGAATTATGAATGTTAC---CGATTCGGCATTTCCCCA------------TCC 1203
                 |||:::         |||         |||:::   :||           |||
Db        337 ------LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSer 354

Qy       1204 ACAAATGCG---------CTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATC 1254
                 |||             :::::::|||        :::      :::|||   :::
Db        355 GlnAsnGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerVal 374

Qy       1255 TTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
                :::|||           ||||||||||||||   :::
Db        375 TyrAspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 7
JC7246
pepsinogen C - common marmoset
C;Species: Callithrix jacchus (common marmoset)
C;Date: 09-Jun-2000 #sequence_revision 09-Jun-2000 #text_change 21-Jul-2000
C;Accession: JC7246
R;Kageyama, T.
J. Biochem. 127, 761-770, 2000
A;Title: New world monkey pepsinogens A and C, and prochymosins. Purification, character
A;Reference number: JC7245
A;Accession: JC7246
A;Molecule type: mRNA
A;Residues: 1-388 <KAG>
A;Cross-references: DDBJ:AB038385
A;Experimental source: strain NW791
C;Comment: This protein, a zymogen of pepsins, is the major proteolytic enzyme in verteb
volution of mammalian orders and families.
C;Superfamily: pepsin
C;Keywords: gastric juice; zymogen

Alignment Scores:
Pred. No.:              2.82e-17        Length:         388
Score:                  351.50          Matches:        112
Percent Similarity:     45.16%          Conservative:   56
Best Local Similarity:  30.11%          Mismatches:     115
Query Match:            10.53%          Indels:         89
DB:                     2               Gaps:           17

US-09-806-194-1 (1-1804) x JC7246 (1-388)

Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
              |:|:::   |||:::   ||||||||||||||||||          :::|||   ||:.||
Db         73 TyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThr 92

Qy        334 GGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCGCAC 389
              ||||||||||:|      |||                    |||   ||    |::
Db         93 GlySerSerAsnLeuTrpValProSerValTyrCysGlnSerGlnAlaCysThrSerHis 112

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
              |||               |||:::         ||||||||||||   |||   |||
Db        113 Ser------------ArgPheAsnProSerAlaSerSerThrTyrSerSerAsnGlyGln 128

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
              :::::::::|||      ||||||   |||||||||   |:|   |:|   :::
Db        129 ThrPheSerLeuGlnTyrGlySerGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 148

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAAICA 549
              |||:::                |||      |||                   |||
Db        149 ThrVal--------------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSer 163

Qy        550 GAGAATTTCTTTTTGCCTGGG--------------ATTAAATGGAATGGAATACTTGGC 594
              ||||||         ||||||                    :::::::::|||||||:::|||
Db        164 GluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 181

Qy        595 CTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTG 654
              |||||||||       |||:::          ::::::      |||             :::
Db        182 LeuAlaTyrProAlaLeuSerMetGlyGlyAlaThr------ThrAlaMetGlnGlyMet 199

Qy        655 GTGACACAAGCAAACATCCCCAAC---GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTG 711
              :::   :::       :::   :::   |||||||||       :::
Db        200 LeuGlnGluGlyAlaLeuThrSerProValPheSerPheTyrLeuSerAsnGln------ 217

Qy        712 CCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGT 771
                        |||:::::::||||||:::::::::   ||||||::::::   |||
Db        218 ------------GlnGlySerSerGlyGlyAlaValIlePheGlyGlyValAspSerSer 233

Qy        772 TTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAA 831
              ||||||   |||   |||:::::::   ||||::   :::|||   |||:::||||||
Db        234 LeuTyrThrGlyGlnIleTyrTrpAlaProValThrGlnGluLeuTyrTrpGlnIleGly 253

Qy        832 ATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCA 891
              |||   :::       ||||||||||||:::             |||   |||
Db        254 IleGluGluPheLeuIleGlyGlyGlnAlaSerGlyTrp---CysSerGlu------Gly 270

Qy        892 GACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGAT 951
                 :::|||||||||||||:::||||||:::||||||   :::|||||||:::
Db        271 CysGlnAlaIleValAspThrGlyThrSerLeuLeuThrValProGlnGlnTyrMetSer 290

Qy        952 GCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGG 1011
              |||   :::|||||||
Db        291 AlaPheLeuGluAla---------------------------------------------- 295

Qy       1012 ACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTC------ 1065
              ||||||:::|||                        |||       :::      |||
Db        296 ThrGlyAlaGln--------------------GluAspGluTyrGlyGlnPheLeuVal 308

Qy       1066 ------------------------CCTAAAATCTCCATCTACCTGAGAGATGAGAACTCC 1101
                                      |||   ::::::      :::
Db        309 AsnCysAspSerIleGlnAsnLeuProThrLeuThrPheIleIle--------------- 323

Qy       1102 AGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCC 1161
                                    :::   |||      ||||||
Db        324 ---AsnGlyValGluPheProLeuProProSerSerTyrIle------------------ 336

Qy       1162 GGCCTGAATTATGAATGTTAC---CGATTCGGCATTTCCCCA------------------ 1200
                 |||:::         :||         |||:::   |||
Db        337 ---LeuSerAsnAsnGlyTyrCysThrValGlyValGluProThrTyrLeuSerSerGln 355

Qy       1201 ---TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
                  |||             ::::::|||         :::      :::|||   :::|||
Db        356 AsnSerGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValPhe 375

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
                 |||           |||'||||||||||||   :::
Db        376 AspLeuGlyAsnAsnArgValGlyPheAlaThrAla 387

RESULT 8
B43356
gastricsin (EC 3.4.23.3) precursor - guinea pig
N;Alternate names: pepsin C
C;Species: Cavia porcellus (guinea pig)
C;Date: 03-Feb-1994 #sequence_revision 03-Feb-1994 #text_change 22-Jun-1999
C;Accession: B43356
R;Kageyama, T.; Ichinose, M.; Tsukada, S.; Miki, K.; Kurokawa, K.; Koiwai, O.; Tanji,
J. Biol. Chem. 267, 16450-16459, 1992
A;Title: Gastric procathepsin E and progastricsin from guinea pig. Purification, mole
A;Reference number: A43356; MUID:92355614; PMID:1644829
A;Accession: B43356
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-394 <KAG>
A;Cross-references: GB:M88652; NID:g191296; PIDN:AAA37053.1; PID:g191297
A;Note: sequence extracted from NCBI backbone (NCBIN:110805, NCBIP:110806)
```

```
 Superfamily: pepsin
 Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach

lignment Scores:
red. No.:              2.45e-15      Length:          394
core:                  324.50        Matches:         107
ercent Similarity:     46.07%        Conservative:    63
est Local Similarity:  29.00%        Mismatches:      116
uery Match:            9.72%         Indels:          83
B:                     2             Gaps:            18

S-09-806-194-1 (1-1804) x B43356 (1-394)

y        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
             |||:::  ::::::  :::|||||||||||||||||     |||:::|||  |||'':|
b         79 TyrPheGlyGlnIleSerLeuGlyThrProProGlnSerPheGlnValLeuPheAspThr 98

y        334 GGAAGCAGTAACTTT----------------------GCCGTGGCAGGAACCCCGCAC 369
             |||||||||||||                      :::::::||  |||  |||
b         99 GlySerSerAsnLeuTrpValProSerValTyrCysSerSerLeuAlaCysThrThrHis 118

y        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
             |||  ||||:::         |||||||||.||  :::
b        119 -----------ThrArgPheAsnProArgAspSerSerThrTyrValAlaThrAspGln 134

y        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
             :::::::::|||     ||||||  ||||||     |||  |||  :::
b        135 SerPheSerLeuGluTyrGlyThrGlySerLeuThrGlyValPheGlyTyrAspThrMet 154

y        490 ACCATC--------------CCCAAA---GGCTTCAATACTTCTTTTCTTGTCAACATT 531
             ||||||              ||||||     |||      |||
b        155 ThrIleGlnAspIleGlnValProLysGlnGluPheGlyLeuSer-------------- 169

y        532 GCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGG--------------ATTAAA 576
             |||:::|||           ||||||                          :::
b        170 -----------GluThrGlu-----------ProGlySerAspPheValTyrAlaGlu 181

y        577 TGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAG 636
             ::::::|||||||||||||||||   |||      |||:::::::      ::::::
b        182 PheAspGlyIleLeuGlyLeuGlyTyrProGlyLeuSerGluGlyGlyAlaThr------ 199

y        637 ACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAG 693
             |||          |||:::  :::      :::      :::::::|||||:::
b        200 ThrAlaMetGlnGlyLeuLeuArgGluGlyAlaLeuSerGlnSerLeuPheSerValTyr 219

y        694 ATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCT------GGGACCAACGGAGGTAGTCTT 747
             :::                  ||||||     |||:::::::   |||   '||
b        220 Leu----------------------GlySerGlnGlnGlySerAspGluGlyGlnLeu 231

y        748 GTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAG 807
             :::||||||||:::::   |||||||||  |||||||||:::::|||||||:::
b        232 IleLeuGlyGlyValAspGluSerLeuTyrThrGlyAspIleTyrTrpThrProValThr 251

Qy       808 GAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAAT 867
             :::|||  |||:::|||||  |||      |||  |||  :::
Db       252 GlnGluLeuTyrTrpGlnIleGlyIleGluGlyPheLeuIleAspGlySerAlaSerGly 271

Qy       868 CTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTG 927
                 |||           :::  ||||||||||:::|||||:::||||||
Db       272 TrpCysSerArg---------GlyCysGlnGlyIleValAspThrGlyThrSerLeuLeu 288

Qy       928 CGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATT 987
                ::|||           :::|||:::|||:::
Db       289 ThrValProSerAspTyrLeuSerThrLeuValGlnAlaIleGlyAlaGluGlu------ 306

Qy       988 CCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAA 1047
             |||:::  :::  |||           :::::|||  :::      :::
Db       307 AsnGluTyrGlyGluTyrPhe--------------ValSerCysSerSerIleGlnAsp 321

Qy      1048 ACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGG 1107
             |||       |||         :::
Db       322 LeuProThrLeuThrPheValIleSerGlyVal-------------------------- 332

Qy      1108 TCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCC--------------ATG 1152
                          :::  |||      ||||||                      |||
Db       333 ------GluPheProLeuSerProSerAlaTyrIleLeuSerGlyGluAsnTyrCysMet 350

Qy      1153 ATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAAT--- 1209
             :::|||          |||                ::::::::
Db       351 ValGlyLeuGluSerThrTyr-----------------ValSerProGlyGlyGlyGlu 364

Qy      1210 ---GCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCC 1266
                    ::::::|||         :::      :::|||   :::::::|||   |||
Db       365 ProValTrpIleLeuGlyAspValPheLeuArgSerTyrTyrSerValTyrAspLeuAla 384

Qy      1267 CAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
                    ||||||||||||||   :::
Db       385 AsnAsnArgValGlyPheAlaThrAla 393

RESULT 9
JC7575
pepsinogen A - bullfrog
C;Species: Rana catesbeiana (bullfrog)
C;Date: 30-Jun-2001 #sequence_revision 30-Jun-2001 #text_change 03-Aug-2001
C;Accession: JC7575
R;Ikuzawa, M.; Inokuchi, T.; Kobayashi, K.; Yasumasu, S.
J. Biochem. 129, 147-153, 2001
A;Title: Amphibian pepsinogens: Purification and characterization of Xenopus pepsinog
A;Reference number: JC7573; MUID:21064922; PMID:11134969
A;Contents: Stomach
A;Accession: JC7575
A;Molecule type: mRNA
A;Residues: 1-385 <IKU>
A;Cross-references: DDBJ:AB045376
C;Comment: This protein is a zymogen for gastric aspartic proteinase, with pepsin-lik
C;Genetics:
A;Gene: PgA
C;Superfamily: pepsin
C;Keywords: stomach; zymogen

Alignment Scores:
Pred. No.:              5.14e-15      Length:          385
Score:                  320.00        Matches:         111
Percent Similarity:     44.61%        Conservative:    67
Best Local Similarity:  27.82%        Mismatches:      147
Query Match:            9.58%         Indels:          74
DB:                     2             Gaps:            15

US-09-806-194-1 (1-1804) x JC7575 (1-385)

Qy       148 GGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCG 207
             |'|        :::::|||          |||          ||||::|||||
Db        39 GlyAspTyrLeuLysLysHisHisTyrAsnProAlaThrLysTyrPheProSerLeuAla 58

Qy       208 ICCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGC 267
                 :::|||                        :::   ||||||
Db        59 GlnAlaSerGly----------------------GluProLeuGlnAsnTyrMetAsp 70

Qy       268 CGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTT 327
                 |||:::      :::   ||||||||||||||||||         ::::::
Db        71 IleGluTyrPheGlyThrIleSerIleGlyThrProProGlnSerPheThrValIlePhe 90

Qy       328 GACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACG--- 384
             |||||||||||||||||      |||         |||      |||      :::
Db        91 AspThrGlySerSerAsnLeuTrpVal---------ProSerValTyrCysSerSerPro 107

Qy       385 -----------------TACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGC 426
                             |||:::  ::::::|||||||||:::::::::
Db       108 AlaCysThrAsnHisHisMetPheAsnProGlnGlnSerSerThrPheGlnAlaThrAsn 127

Qy       427 TTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTC 486
                 |||:::::::::|||      ||||||   :::|||||:::|||   |||
Db       128 ThrProValSerIleGlnTyrGlyThrGlySerMetSerGlyPheLeuGlyTyrAspThr 147

Qy       487 GTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAA 546
             |||  :::           |||                 |||      :::  :::
```

```
Db        148 ValGlnValGly---------AsnIleGlnIleThrAsnGlnIlePheGlyLeuSerGln 184

Qy        547 TCAGAG---AATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTAT 603

Db        155 SerGluProGlySerPheLeuTyrTyrSerProPheAspGlyIleLeuGlyLeuAlaPhe 184

Qy        604 GCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCITCTTCGACTCCCTGGTGACACAA 663

Db        185 ProSerLeuAla------SerSerGlnAlaThrProValPheAspAsnMetTrpAsnGln 202

Qy        664 GCAAACATCCCC---AACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTIGCCCGTTGCT 720

Db        203 GlyLeuIleProGlnAspLeuPheSerValTyrLeu----------Ser 215

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780

Db        216 SerGlnGlyGlnSerGlySerPheValLeuPheGlyGlyValAspThrSerTyrTyrThr 235

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840

Db        236 GlyAsnLeuAsnTrpValProLeuThrAlaGluThrTyrTrpGlnIleThrValAspSer 255

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900

Db        256 IleSerIleGlyGlyGlnValIleAlaCysSer----------GlySerCysSerAla 271

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960

Db        272 IleValAspThrGlyThrSerLeuLeuAlaGlyProSerThrProIle---AlaAsnIle 290

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020

Db        291 GlnTyrTyrIleGlyAlaAsn---------GlnAspSerAsnGlyGlnTyrVal------ 305

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT------------ 1068

Db        306 ---IleAsnCysAsnAsnIleSerAsnMetProThrValValPheThrIleAsnGlyVal 324

Qy       1069 ------------AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGC------AGGTCA 1110

Db        325 GlnTyrProLeuProAlaSerAlaTyrValArgGlnSerGlnGlnSerCysThrSerGly 344

Qy       1111 TTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAAT 1170

Db        345 PheGlnAlaMetAsnLeuPro---------------------------------------- 351

Qy       1171 TATGAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACG 1230

Db        352 ----------------------ThrSerSerGlyAspLeuTrpIleLeuGlyAspVal 363

Qy       1231 GTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCA 1287

Db        364 PheIleArgGluTyrTyrValValPheAspArgAlaAsnAsnTyrValAlaMetAla 382

RESULT 10
REMSK
renin (EC 3.4.23.15) precursor, renal - mouse
N;Alternate names: angiotensin-forming enzyme; angiotensinogenase; renin 1
C;Species: Mus musculus (house mouse)
C;Date: 30-Jun-1987 #sequence_revision 30-Jun-1987 #text_change 18-Jun-1993
C;Accession: A00989; S07636; A22766; A22058; I57576; A05137; JH0083
R;Holm, I.; Ollo, R.; Panthier, J.J.; Rougeon, F.
EMBO J. 3, 557-562, 1984
A;Title: Evolution of aspartyl proteases by gene duplication: the mouse renin gene is or
A;Reference number: A00989; MUID:84182525; PMID:6370686
A;Accession: A00989
A;Molecule type: DNA
A;Residues: 1-402 <HOL>
A;Cross-references: EMBL:X00850
R;Kim, W.S.; Murakami, K.; Nakayama, K.
Nucleic Acids Res. 17, 9480, 1989
A;Title: Nucleotide sequence of a cDNA coding for mouse Renl preprorenin.
A;Reference number: S07636; MUID:90067953; PMID:2685761
A;Accession: S07636
A;Molecule type: mRNA
A;Residues: 1-402 <KIM>
A;Cross-references: EMBL:X16642; NID:g53930; PIDN:CAA34636.1; PID:g53931
R;Mullins, J.J.; Burt, D.W.; Windass, J.D.; McTurk, P.; George, H.; Brammar, W.J.
EMBO J. 1, 1461-1466, 1982
A;Title: Molecular cloning of two distinct renin genes from the DBA/2 mouse.
A;Reference number: A90968; MUID:84207899; PMID:6327270
A;Accession: A22766
A;Molecule type: mRNA
A;Residues: 269-314,'D',316 <MUL>
R;Panthier, J.J.; Dreyfus, M.; Roux, D.T.L.; Rougeon, F.
Proc. Natl. Acad. Sci. U.S.A. 81, 5489-5493, 1984
A;Title: Mouse kidney and submaxillary gland renin genes differ in their 5' putative
A;Reference number: A22058; MUID:84298161; PMID:6089205
A;Accession: A22058
A;Molecule type: DNA
A;Residues: 1-30 <PAN>
R;Field, L.J.; Philbrick, W.M.; Howles, P.N.; Dickinson, D.P.; McGowan, R.A.; Gross,
Mol. Cell. Biol. 4, 2321-2331, 1984
A;Title: Expression of tissue-specific Ren-1 and Ren-2 genes of mice: Comparative ana
A;Reference number: I57576; MUID:85085936; PMID:6392850
A;Accession: I57576
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-31 <RES>
A;Cross-references: GB:K02800; NID:g200689; PIDN:AAA40044.1; PID:g200690
C;Comment: The only known function of renal renin is to release angiotensin I from an
creased sodium retention by the kidney.
C;Comment: Renal renin is synthesized by the juxtaglomerular cells of the kidney in r
C;Genetics:
A;Gene: Ren-1
A;Introns: 31/2; 81/3; 123/1; 162/3; 228/2; 268/2; 316/3; 349/3
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; blood pressure control; glycoprotein; hydrolase; kid
F;1-21/Domain: signal sequence #status predicted <SIG>
F;22-64/Domain: propeptide #status predicted <PRO>
F;65-402/Product: renin #status predicted <MAT>
F;69,139,320/Binding site: carbohydrate (Asn) (covalent) #status predicted
F;102,287/Active site: Asp #status predicted

Alignment Scores:
Pred. No.:                5.18e-15        Length:          402
Score:                    320.00          Matches:         126
Percent Similarity:       43.54%          Conservative:    66
Best Local Similarity:    28.57%          Mismatches:      181
Query Match:              9.58%           Indels:          68
DB:                       1               Gaps:            21

US-09-806-194-1 (1-1804) x REMSK (1-402)

Qy         28 CTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCCTTC 87

Db          6 MetProLeuTrpAlaLeuLeuLeuLeu--------------TrpSerProCysThrPhe 20

Qy         88 ACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGA--- 144

Db         21 SerLeuProThrArgThrAlaThrPheGluArgIleProLeuLysLysMetProSerVal 40

Qy        145 CCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCCCTGGAG--------- 195

Db         41 ArgGluIleLeuGluArgGlyValAspMetThrArgLeuSerAlaGluTrpGlyVal 60

Qy        196 ------------CCTGCC---------CTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTG 234

Db         61 PheThrLysArgProSerLeuThrAsnLeuThrSerProValValLeuThrAsnTyrLeu 80

Qy        235 GCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATC 294

Db         81 ------------AsnThrGln-----------------TyrTyrGlyGluIleGlyIle 90

Qy        295 GGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTG 354
```

```
b    91 GlyThrProProGlnThrPheLysValIlePheAspThrGlySerAlaAsnLeuTrpVal 110

y   355 GCAGGAACCCCGCACTCCTAC------------------ATAGACACGTACTTTGACACA 396
b   111 ProSerThrLysCysSerArgLeuTyrLeuAlaCysGlyIleHisSerLeuTyrGluSer 130

y   397 GAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGA 456
b   131 SerAspSerSerSerTyrMetGluAsnGlySerAspPheThrIleHisTyrGlySerGly 150

y   457 AGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCT 516
b   151 ArgValLysGlyPheLeuSerGlnAspSerValThrVal---GlyGlyIleThrValThr 169

y   517 TTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAA 576
b   170 GlnThrPheGlyGluValThrGluLeuProLeuIleProPheMetLeu------AlaLys 187

y   577 TGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAG 636
b   188 PheAspGlyValLeuGlyMetGlyPhePro------AlaGlnAlaValGlyGlyValThr 205

y   637 ACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTCTCCATGCAG 693
b   206 ProValPheAspHisIleLeuSerGlnGlyValLeuLysGluGluValPheSerValTyr 225

y   694 ATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTG 753
b   226 Tyr------------------AsnArgGlySerHisLeuLeuGlyGlyGluValValLeu 239

y   754 GGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAG 813
b   240 GlyGlySerAspProGlnHisTyrGlnGlyAsnPheHisTyrValSerIleSerLysThr 259

y   814 TGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGAC 873
b   260 AspSerTrpGlnIleThrMetLysGlyValSerValGly------SerSerThrLeuLeu 277

y   874 TGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTG 933
b   278 CysGluGluGlyCysAla------ValValValAspThrGlySerSerPheIleSerAla 295

y   934 CCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTG---GCCCGCGCATCTCTGATTCCA 990
b   296 ProThrSerSerLeuLysLeuIleMetGlnAlaLeuGlyAlaLysGluLysArgIleGlu 315

y   991 GAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACA 1050
b   316 GluTyr-----------------------ValValAsnCys---------SerGlnVal 324

y  1051 CCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCA 1110
b   325 Pro------ThrLeuProAspIleSerPheAspLeu------------GlyGlyArgAla 338

y  1111 TTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTG--- 1167
b   339 TyrThrLeuSerSerThrAspTyrValLeuGlnTyrProAsnArgArgAspLysLeuCys 358

y  1168 AATTATGAATGTTACCGATTCGGCATTTCCCCATCCACAAATGCGCTG---GTGATCGGT 1224
b   359 ThrLeuAlaLeuHisAlaMetAspIleProProProThrGlyProValTrpValLeuGly 378

y  1225 GCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTC 1284
b   379 AlaThrPheIleArgLysPheTyrThrGluPheAspArgHisAsnAsnArgIleGlyPhe 398

y  1285 GCA 1287
b   399 Ala 399
```

RESULT 11
S66516
oryzasin (EC 3.4.23.-) precursor - rice

N;Alternate names: aspartic proteinase 1
C;Species: Oryza sativa (rice)
C;Date: 28-Oct-1996 #sequence_revision 13-Mar-1997 #text_change 20-Jun-2000
C;Accession: S66516; S66517
R;Asakura, T.; Watanabe, H.; Abe, K.; Arai, S.
Eur. J. Biochem. 232, 77-83, 1995
A;Title: Rice aspartic proteinase, oryzasin, expressed during seed ripening and germi
A;Reference number: S66516; MUID:96048031; PMID:7556174
A;Accession: S66516
A;*Molecule type: DNA*
A;Residues: 1-509 <ASA>
A;Cross-references: EMBL:D32165; NID:g511665; PIDN:BAA06876.1; PID:g1030715
A;Accession: S66517
A;Molecule type: mRNA
A;Residues: 1-509 <AS2>
A;Cross-references: EMBL:D32144; NID:g1255684; PIDN:BAA06875.1; PID:g1711289
C;Comment: The pair of saposin repeat homology domains tagged SAP1 and SAP2 represent
C;Genetics:
A;Introns: 119/3; 140/1; 171/3; 209/2; 265/3; 279/1; 300/3; 338/3; 360/2; 412/3; 452/
C;*Superfamily: oryzasin; saposin repeat homology*
C;Keywords: aspartic proteinase; hydrolase
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-68/Domain: propeptide #status predicted <PRO>
F;68-509/Product: aspartic proteinase 1 #status predicted <MAT>
F;316-361/Domain: saposin repeat homology #status atypical <SAP1>
F;370-420/Domain: saposin repeat homology #status atypical <SAP2>
F;103,290/Active site: Asp #status predicted

Alignment Scores:

Pred. No.:	1.58e-14	*Length:*	509
Score:	313.50	Matches:	127
Percent Similarity:	36.59%	Conservative:	75
Best Local Similarity:	23.01%	Mismatches:	179
Query Match:	9.39%	Indels:	171
DB:	2	Gaps:	19

US-09-806-194-1 (1-1804) x S66516 (1-509)

```
Qy     7 GCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCG 66
Db     5 SerValAlaLeuValLeuLeuAlaAlaValLeuLeuGlnAlaLeuLeuProAlaSerAla 24

Qy    67 GAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTA 126
Db    25 GluGluGlyLeuValArgIleAlaLeuLysLysArgProIleAspGluAsnSerArgVal 44

Qy   127 GTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTC 186
Db    45 AlaAlaArgLeuSerGly---------------GluGluGlyAlaArgArgLeuGlyLeu 59

Qy   187 GCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGAC 246
Db    60 ArgGlyAlaAsnSerLeuGlyGlyGlyGlyGlyGluGlyAspIleValAlaLeuLysAsn 79

Qy   247 AACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCG 306
Db    80 TyrMetAsnAlaGln------------TyrPheGlyGluIleGlyValGlyThrProPro 95

Qy   307 CAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAAC--------------------- 345
Db    96 GlnLysPheThrValIlePheAspThrGlySerSerAsnLeuTrpValProSerAlaLys 115

Qy   346 ------TTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAG 399
Db   116 CysTyrPheSerIleAlaCysPhePheHisSer------------ArgTyrLysSerGly 131

Qy   400 AGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGC 459
Db   132 GlnSerSerThrTyrGlnLysAsnGlyLysProAlaAlaIleGlnTyrGlyThrGlySer 151

Qy   460 TGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTT 519
Db   152 IleAlaGlyPhePheSerGluAspSerValThrValGlyAsp------------------ 165
```

```
Qy       520 CTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTT-----------TTGCCT 567
Db       166 --------------LeuValValLysAspGlnGluPheIleGluAlaThrLysGluPro 180

Qy       568 GGGATT--------------AAATGGAATGGAATACTTGGCCTAGCTTATGCCAGACTT 612
Db       181 GlyLeuThrPheMetValAlaLysPheAspGlyIleLeuGlyLeuGlyPheGlnGluIle 200

Qy       613 GCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC 672
Db       201 SerValGlyAspAla----------------------------------------Val 206

Qy       673 CCCAACGTTTTCTCCATGCAGATGTGTGGA---GCCGGCTTGCCCGTTGCTGGATCTGGG 729
Db       207 ProValTrpTyrLysMetValGluGlnGlyLeuValSerGluProValPheSerPheTrp 226

Qy       730 ACCAAC--------------------GGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db       227 PheAsnArgHisSerAspGluGlyGluGlyGlyGluIleValPheGlyGlyMetAspPro 246

Qy       769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db       247 SerHisTyrLysGlyAsnHisThrTyrValProValSerGlnLysGlyTyrTrpGlnPhe 266

Qy       829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db       267 GluMetGlyAspValLeuIleGlyGlyLysThrThrGlyPhe---CysAla------Ser 283

Qy       889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db       284 GlyCysSerAlaIleAlaAspSerGlyThrSerLeuLeuAlaGlyProThrAlaIleIle 303

Qy       949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAA--------------- 993
Db       304 ThrGluIleAsnGluLysIleGlyAlaThrGlyValValSerGlnGluCysLysThrVal 323

Qy       993 ------------------------------------------------------------ 993
Db       324 ValSerGlnTyrGlyGlnGlnIleLeuAspLeuLeuLeuAlaGluThrGlnProSerLys 343

Qy       993 ------------------------------------------------------------ 993
Db       344 IleCysSerGlnValGlyLeuCysThrPheAspGlyLysHisGlyValSerAlaGlyIle 363

Qy       994 --------------------------TTCTCTGATGGTTTC------------ ------ 1008
Db       364 LysSerValValAspAspGluAlaGlyGluSerAsnGlyLeuGlnSerGlyProMetCys 383

Qy      1009 ----------------------TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCG 1044
Db       384 AsnAlaCysGluMetAlaValValTrpMetGlnAsnGlnLeuAlaGlnAsnLysThrGln 403

Qy      1045 GAAACACCTTGGTCTTAC------------------------------------------ 1062
Db       404 AspLeuIleLeuAsnTyrIleAsnGlnLeuCysAspLysLeuProSerProMetGlyGlu 423

Qy      1063 ----------------------------TTCCCTAAAATCTCCATCTACCTGAGAGAT 1092
Db       424 SerSerValAspCysGlySerLeuAlaSerMetProGluIleSerPheThrIleGlyAla 443

Qy      1093 GAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATG 1152
Db       444 Lys-----------------LysPheAlaLeuLysProGluGluTyrIleLeu---Lys 456

Qy      1153 ATGGGGGCCGGCCTGAATTATGAATGTTAC-----------CGATTCGGCATTTCCCCA 1200
Db       457 ValGlyGluGlyAlaAlaAlaGlnCysIleSerGlyPheThrAlaMetAspIleProPro 476

Qy      1201 TCCACAAATGCGCTG---GTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
Db       477 ProArgGlyProLeuTrpIleLeuGlyAspValPheMetGlyAlaTyrHisThrValPhe 496
```

```
Qy      1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293
Db       497 AspTyrGlyLysMetArgValGlyPheAlaLysSer 508

RESULT 12
A24608
gastricsin (EC 3.4.23.3) precursor - rat
N;Alternate names: pepsinogen C
N;Contains: pepsin A (EC 3.4.23.1) precursor
C;Species: Rattus norvegicus (Norway rat)
C;Date: 30-Jun-1988 #sequence_revision 05-Aug-1994 #text_change 18-Jun-1999
C;Accession: A33510; A24608; C22434; A05145; A61298
R;Ishihara, T.; Ichihara, Y.; Hayano, T.; Katsura, I.; Sogawa, K.; Fujii-Kuriyama, Y.
J. Biol. Chem. 264, 10193-10199, 1989
A;Title: Primary structure and transcriptional regulation of rat pepsinogen C gene.
A;Reference number: A33510; MUID:89255508; PMID:2722863
A;Accession: A33510
A;Molecule type: DNA
A;Residues: 1-392 <ISH>
A;Cross-references: GB:M25985
R;Ichihara, Y.; Sogawa, K.; Morohashi, K.; Fujii-Kuriyama, Y.; Takahashi, K.
Eur. J. Biochem. 161, 7-12, 1986
A;Title: Nucleotide sequence of a nearly full-length cDNA coding for pepsinogen of ra
A;Reference number: A24608; MUID:87054020; PMID:3780741
A;Accession: A24608
A;Molecule type: mRNA
A;Residues: 1-392 <ICH>
A;Cross-references: GB:X04644; NID:g56880; PIDN:CAA28305.1; PID:g56881
R;Ichihara, Y.; Sogawa, K.; Takahashi, K.
J. Biochem. 98, 483-492, 1985
A;Title: Isolation of human, swine, and rat prepepsinogens and calf preprochymosin, a
A;Reference number: A22434; MUID:86059312; PMID:2415509
A;Accession: C22434
A;Molecule type: protein
A;Residues: 1-19,'X',21-23,'X',25-29 <IC2>
R;Arai, K.M.; Muto, N.; Tani, S.; Akahane, K.
Biochim. Biophys. Acta 788, 256-261, 1984
A;Title: The N-terminal sequence of rat pepsinogen.
A;Reference number: A05145; MUID:84257697; PMID:6743670
A;Accession: A05145
A;Molecule type: protein
A;Residues: 17-30,'Q',32-102,'A',104-108,'L',110-112 <ARA>
A;Experimental source: Wistar strain
R;Ichihara, Y.; Sogawa, K.; Takahashi, K.
J. Biochem. 92, 603-606, 1982
A;Title: Rat gastric prepepsinogen: in vitro synthesis and partial amino-terminal sig
A;Reference number: A61298; MUID:83030750; PMID:6182139
A;Accession: A61298
A;Molecule type: protein
A;Residues: 1,'XX',4-6,'X',8-9,'X',11,'X',13-14,'XXX',18-19,'X',21,'X',23,'XX',26,'X'
C;Comment: This enzyme has more restricted specificity than pepsin A. It is the major
C;Genetics:
A;Introns: 20/2; 73/3; 113/1; 152/3; 219/2; 259/2; 309/3; 342/3
A;Note: there are at least two very similar genes for gastricsin in rat
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; gastric juice; hydrolase; protein digestion; stomach
F;1-16/Domain: signal sequence #status experimental <SIG>
F;17-392/Product: pepsinogen #status experimental <MAT>
F;17-62/Domain: activation peptide #status experimental <ACT>
F;94,280/Active site: Asp #status predicted
F;107-112,270-275,314-347/Disulfide bonds: #status predicted

Alignment Scores:
Pred. No.:                1.64e-14        Length:          392
Score:                    313.00          Matches:         105
Percent Similarity:       45.22%          Conservative:    56
Best Local Similarity:    29.49%          Mismatches:      139
Query Match:              9.37%           Indels:          56
DB:                       1               Gaps:            16

US-09-806-194-1 (1-1804) x A24608 (1-392)

Qy        274 TACTACCTGGAGATGCTGATCGGGACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
```

```
b         76 TyrPheGlyGluIleSerIleGlyThrProProGlnAsnPheLeuValLeuPheAspThr 95

y        334 GGAAGCAGTAACTTTGCCGTG----------------------GCAGGAACCCCGCAC 369

b         96 GlySerSerAsnLeuTrpValSerSerValTyrCysGlnSerGluAlaCysThrThrHis 115

y        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429

b        116 Ala------------ArgPheAsnProSerLysSerSerThrTyrTyrThrGluGlyGln 131

y        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489

b        132 ThrPheSerLeuGlnTyrGlyThrGlySerLeuThrGlyPhePheGlyTyrAspThrLeu 151

y        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549

b        152 ThrVal--------------GlnSerIleGlnValProAsnGlnGluPheGlyLeuSer 166

y        550 GAGAATTTCTTTTTGCCTGGG--------------ATTAAATGGAATGGAATACTTGGC 594

b        167 GluAsn------GluProGlyThrAsnPheValTyrAlaGlnPheAspGlyIleMetGly 184

y        595 CTAGCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTGCCTG 654

b        185 LeuAlaTyrProGlyLeuSer------SerGlyGlyAlaThrThrAlaLeuGlnGlyMet 202

y        655 GTGACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCC 714

b        203 LeuGlyGlu-----------GlyAlaLeuSerGlnProLeuPheGlyValTyrLeu--- 217

y        715 GTTGCTGGATCT------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768

b        218 ------GlySerGlnGlnGlySerAsnGlyGlyGlnIleValPheGlyGlyValAspLys 235

y        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828

b        236 AsnLeuTyrThrGlyGluIleThrTrpValProValThrGlnGluLeuTyrTrpGlnIle 255

y        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888

b        256 ThrIleAspAspPheLeuIleGlyAspGlnAlaSerGlyTrp---CysSerSerGlnGly 274

y        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCCCTGCCCCAGAAGGTGTTT 948

b        275 Cys---GlnGlyIleValAspThrGlyThrSerLeuLeuValMetProAlaGlnTyrLeu 293

y        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008

b        294 SerGluLeuLeuGlnThrIleGlyAlaGlnGlu------GlyGluTyrGlyGluTyrPhe 311

y       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068

b        312 ---------------ValSerCysAspSerValSerSer---------------LeuPro 321

y       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128

b        322 ThrLeuSerPheValLeu-----------------AsnGlyValGlnPheProLeuSer 335

y       1129 CCTCAGCTTTAC---ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGA 1185

b        336 ProSerSerTyrIleIleGlnGluAspAsnPheCysMetValGlyLeuGluSerIleSer 355

y       1186 ITCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245

b        356 LeuThrSerGluSerGlyGlnProLeuTrpIleLeuGlyAspValPheLeuArgSerTyr 375

y       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGC 1293

b        376 TyrAlaIlePheAspMetGlyAsnAsnLysValGlyLeuAlaThrSer 391

RESULT 13
A41443
pepsin (EC 3.4.23.-) precursor, embryonic - chicken
C;Species: Gallus gallus (chicken)
C;Date: 05-Jun-1992 #sequence_revision 05-Jun-1992 #text_change 21-Jul-2000
C;Accession: A41443
R;Hayashi, K.; Agata, K.; Mochii, M.; Yasugi, S.; Eguchi, G.; Mizuno, T.
J. Biochem. 103, 290-296, 1988
A;Title: Molecular cloning and the nucleotide sequence of cDNA for embryonic chicken
A;Reference number: A41443; MUID:88227903; PMID:3131317
A;Accession: A41443
A;Status: preliminary
A;Molecule type: mRNA
A;Residues: 1-383 <HAY>
A;Cross-references: GB:D00215; NID:g2760810; PIDN:BAA00153.1; PID:g222853
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; hydrolase; protein digestion

Alignment Scores:
Pred. No.:              2.68e-14        Length:         383
Score:                  310.00          Matches:        106
Percent Similarity:     42.78%          Conservative:   63
Best Local Similarity:  26.84%          Mismatches:     136
Query Match:            9.28%           Indels:         90
DB:                     2               Gaps:           15

US-09-806-194-1 (1-1804) x A41443 (1-383)

Qy       166 CACGCC------GACGGCTTGGCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGC 219

Db        55 HisAlaProAspValLeuThrValValThrGluProLeuLeu-------------- 69

Qy       220 GCCGCCAACTCTCTTGGCCATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTAC 279

Db        70 ------AsnThrLeuAspMet----------------------------GluTyrTyr 77

Qy       280 CTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGC 339

Db        78 GlyThrIleSerIleGlyThrProProGlnAspPheThrValValPheAspThrGlySer 97

Qy       340 AGTAACTTTGCCGTGGCAGGA-----------ACCCCGCACTCCTACATAGACACGTAC 387

Db        98 SerAsnLeuTrpValProSerValSerCysThrSerProAlaCysGlnSerHisGlnMet 117

Qy       388 TTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTAC 447

Db       118 PheAsnProSerGlnSerSerThrTyrLysSerThrGlyGlnAsnLeuSerIleHisTyr 137

Qy       448 ACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATCCCCAAAGGCTTC 507

Db       138 GlyThrGlyAspMetGluGlyThrValGlyCysAspThrValThrValAlaSerLeuMet 157

Qy       508 AATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCT 567

Db       158 AspThrAsnGlnLeuPheGlyLeuSerThr---SerGluProGlyGlnPhePheVal--- 175

Qy       568 GGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCATCAAGT 627

Db       176 TyrValLysPheAspGlyIleLeuGlyLeuGlyTyrProSerLeuAla-----AlaAsp 193

Qy       628 TCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAACGTTTTC 684

Db       194 GlyIleThrProValPheAspAsnMetValAsnGluSerLeuLeuGluGlnAsnLeuPhe 213

Qy       685 TCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGAGGTAGT 744

Db       214 SerValTyrLeuSerArgGluProMet----------------------GlySerMet 225

Qy       745 CTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACCCCTATT 804

Db       226 ValValPheGlyGlyIleAspGluSerTyrPheThrGlySerIleAsnTrpIleProVal 245

Qy       805 AAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTT 864

Db       246 SerTyrGlnGlyTyrTrpGlnIleSerMetAspSerIleIleValAsnLysGlnGluIle 265
```

```
Qy        865 AATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACCACGCTG 924
                          :::      :::|||||:::|| :::;||||||:::||
Db        266 AlaCysSer-----------SerGlyCysGlnAlaIleIleAspThrGlyThrSerLeu 281

Qy        925 CTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTG 984
              :::      !||           :::   :::      ;|||||;
Db        282 ValAlaGlyProAlaSerAspIleAsnAspIleGlnSerAlaValGly------------ 297

Qy        985 ATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCG 1044
                                                                    |||
Db        298 ----------------------------------------------------AlaAsnGln 300

Qy       1045 GAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAACTCCAGC 1104
                |||        :||                                  |||
Db        301 AsnThrTyrGlyGluTyr-----------------------------------SerVal 318

Qy       1105 AGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGC 1164
                        |||      :::|||   :::            ::::::||;   |||
Db        309 AsnCysSerHisIleLeuAlaMetProAspValValPhe------ValIleGly---Gly 325

Qy       1165 CTGAATTAT-------------------------------------------GAATGTTAC 1182
              :::   |||                                           :||
Db        326 IleGlnTyrProValProAlaLeuAlaTyrThrGluGlnAsnGlyGlnGlyThrCysMet 345

Qy       1183 CGATTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGC 1242
                        :::   |||   :::      ::::::|:|        :::
Db        346 SerSerPheGlnAsnSerSerAlaAspLeuTrpIleLeuGlyAspValPheIleArgVal 365

Qy       1243 TTCTACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCA 1287
              :::|||   |||||||||||||||      ||||||||   |||
Db        366 TyrTyrSerIlePheAspArgAlaAsnAsnArgValGlyLeuAla 380
```

RESULT 14
KHHUD
cathepsin D (EC 3.4.23.5) precursor [validated] - human
N;Alternate names: preprocathepsin D
C;Species: Homo sapiens (man)
C;Date: 28-Dec-1987 #sequence_revision 28-Dec-1987 #text_change 15-Sep-2000
C;Accession: A25771; S30749; PC2066; I59236; I57716
R;Faust, P.L.; Kornfeld, S.; Chirgwin, J.M.
Proc. Natl. Acad. Sci. U.S.A. 82, 4910-4914, 1985
A;Title: Cloning and sequence analysis of cDNA for human cathepsin D.
A;Reference number: A25771; MUID:85270436; PMID:3927292
A;Accession: A25771
A;Molecule type: mRNA
A;Residues: 1-412 <FAU>
A;Cross-references: EMBL:M11233; NID:g181179; PIDN:AAB59529.1; PID:g181180
R;Westley, B.R.; May, F.E.B.
Nucleic Acids Res. 15, 3773-3786, 1987
A;Title: Oestrogen regulates cathepsin D mRNA levels in oestrogen responsive human breas
A;Reference number: S30749; MUID:87231068; PMID:3588310
A;Accession: S30749
A;Molecule type: mRNA
A;Residues: 1-412 <WES>
A;Cross-references: EMBL:X05344; NID:g29677; PIDN:CAA28955.1; PID:g29678
R;May, F.E.B.; Smith, D.J.; Westley, B.R.
Gene 134, 277-282, 1993
A;Title: The human cathepsin D-encoding gene is transcribed from an estrogen-regulated a
A;Reference number: PC2066; MUID:94085791; PMID:8262386
A;Accession: PC2066
A;Molecule type: DNA
A;Residues: 1-23 <MAY>
A;Cross-references: GB:L12980; NID:g291930; PIDN:AAA16314.1; PID:g455429
A;Experimental source: MCF-7 cell
R;Cavailles, V.; Augereau, P.; Rochefort, H.
Proc. Natl. Acad. Sci. U.S.A. 90, 203-207, 1993
A;Title: Cathepsin D gene is controlled by a mixed promoter, and estrogens stimulate onl
A;Reference number: I59236; MUID:93126342; PMID:8419924
A;Accession: I59236
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-22 <CAV1>
A;Cross-references: GB:S52557; NID:g263124; PIDN:AAD13868.1; PID:g4261568
R;Augereau, P.; Miralles, F.; Cavailles, V.; Gaudelet, C.; Parker, M.; Rochefort, H.
Mol. Endocrinol. 8, 693-703, 1994
A;Title: Characterization of the proximal estrogen-responsive element of human cathep
A;Reference number: I57716; MUID:95021301; PMID:7935485
A;Accession: I57716
A;Status: translation not shown; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-22 <CAV2>
A;Cross-references: GB:S74689; NID:g786350; PIDN:AAD14156.1; PID:g4261856
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51839; PDB:1LYA
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, residues 65-161;170-241
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Erickson, J.W.
submitted to the Brookhaven Protein Data Bank, April 1993
A;Reference number: A51840; PDB:1LYB
A;Contents: annotation; X-ray crystallography, 2.5 angstroms, with inhibitor residues
R;Baldwin, E.T.; Bhat, T.N.; Gulnik, S.; Hosur, M.V.; Sowder II, R.C.; Cachau, R.E.;
Proc. Natl. Acad. Sci. U.S.A. 90, 6796-6800, 1993
A;Title: Crystal structures of native and inhibited forms of human cathepsin D: impli
A;Reference number: A48229; MUID:93342076; PMID:8393577
A;Contents: annotation; X-ray crystallography, 2.5 angstroms
C;Comment: Cathepsin D is a ubiquitous lysosomal proteinase.
C;Comment: In addition to the propeptide, residues 163-168 and 411-412 are proteolyti
C;Comment: The carbohydrate bound to 134-Asn contains a mannose-6-phosphate that is b
C;Genetics:
A;Gene: GDB:CTSD
A;Cross-references: GDB:120512; OMIM:116840
A;Map position: 11p15.5-11p15.5
C;Function:
A;Description: limited specificity endopeptidase
A;Pathway: intracellular protein degradation
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradati
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-64/Domain: propeptide #status predicted <PRO>
F;65-162,169-410/Product: cathepsin D #status experimental <MAT>
F;267,329-356/Region: phosphotransferase recognition
F;91-160,110-117,286-290,329-366/Disulfide bonds: #status experimental
F;97,295/Active site: Asp #status experimental
F;134,263/Binding site: carbohydrate (Asn) (covalent) #status experimental

Alignment Scores:

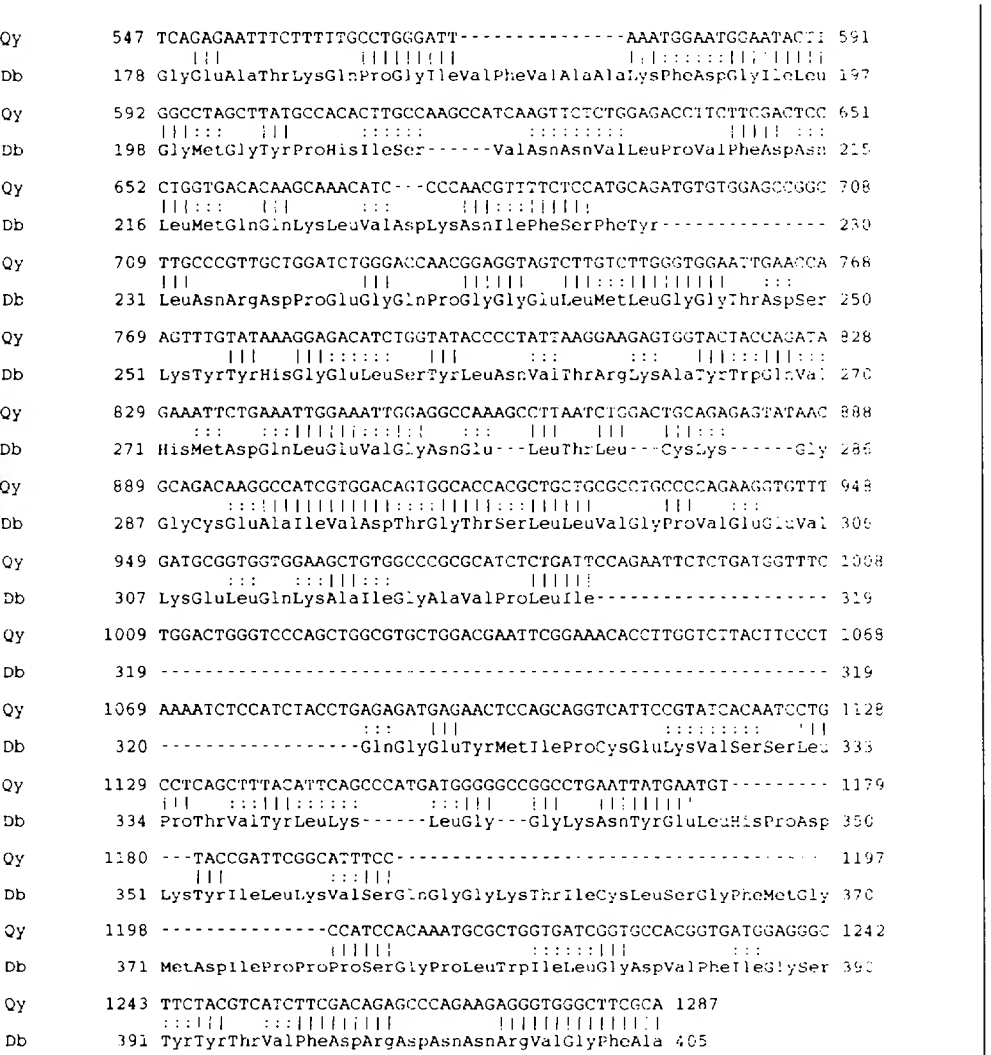
Pred. No.:	3.48e-14	Length:	412
Score:	308.50	Matches:	121
Percent Similarity:	43.85%	Conservative:	75
Best Local Similarity:	27.07%	Mismatches:	180
Query Match:	9.24%	Indels:	71
DB:	1	Gaps:	22

US-09-806-194-1 (1-1804) x KHHUD (1-412)

```
Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
              ||||||||||||                ||| |||||||||      ||| |||
Db          6 LeuLeuProLeuAlaLeu------CysLeuLeuAlaAlaPro------AlaSerAlaLeu 21

Qy         85 TTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGA 144
                    :::|||||||           :::      |||   ::::::          ||
Db         22 ValArgIleProLeuHisLysPheThrSerIleArgArgThrMetSerGluValGlyGly 41

Qy        145 ------------------CCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTC 186
                                  |||       :::::::::       :::
Db         42 SerValGluAspLeuIleAlaLysGlyProValSerLysTyrSerGlnAlaValProAla 61

Qy        187 GCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGAC 246
                 |||      :::      |||           |||::::::
Db         62 ValThrGluGlyProIle------ProGluValLeuLysAsnTyrMet------------ 75

Qy        247 AACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCG 306
                                    ||||||   |||:::   ||||||||||||||||
Db         76 ------------------AspAlaGlnTyrTyrGlyGluIleGlyIleGlyThrProPro 89
```

```
y       307 CAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCG 366
b        90 GlnCysPheThrValValPheAspThrGlySerSerAsnLeuTrpValProSerIleHis 109

y       367 CACTCCTACATAGACACG----------------TACTTTGACACAGAGAGGTCTAGC 408
b       110 CysLysLeuLeuAspIleAlaCysTrpIleHisHisLysTyrAsnSerAspLysSerSer 129

y       409 ACATACCGCTCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGC 468
)b      130 ThrTyrValLysAsnGlyThrSerPheAspIleHisTyrGlySerGlySerLeuSerGly 149

)y      469 TTCGTTGGGGAAGACCTCGTCACCATCCCC------AAAGGCTTCAATACTTCTTTTCTT 522
)b      150 TyrLeuSerGlnAspThrValSerValProCysGlnSerAlaSerSerAlaSerAlaLeu 169

)y      523 ------GTCAACATTGCCACTATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATT--- 573
b       170 GlyGlyValLysValGluArgGlnValPheGlyGluAlaThrLysGlnProGlyIleThr 189

y       574 ------------AAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGCCAAGCCA 621
b       190 PheIleAlaAlaLysPheAspGlyIleLeuGlyMetAlaTyrProArgIleSer------ 207

y       622 TCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATC---CCCAAC 678
)b      208 ValAsnAsnValLeuProValPheAspAsnLeuMetGlnGlnLysLeuValAspGlnAsn 227

)y      679 GTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCTGGATCTGGGACCAACGGA 738
)b      228 IlePheSerPheTyr--------------LeuSerArgAspProAspAlaGlnProGly 242

)y      739 CGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGACATCTGGTATACC 798
)b      243 GlyGluLeuMetLeuGlyGlyThrAspSerLysTyrTyrLysGlySerLeuSerTyrLeu 262

)y      799 CCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGAAATTGGAGGCCAA 858
)b      263 AsnValThrArgLysAlaTyrTrpGlnValHisLeuAspGlnValGluVal---AlaSer 281

)y      859 AGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGTGGACAGTGGCACC 918
)b      282 GlyLeuThrLeu---CysLysGlu------GlyCysGluAlaIleValAspThrGlyThr 298

)y      919 ACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGCTGTGGCCCGCGGCA 978
)b      299 SerLeuMetValGlyPro---------ValAspGluValArgGluLeuGlnLysAlaIle 315

)y      979 TCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCTGGCGTGCTGGACG 1038
Db      316 GlyAlaValProLeuIleGlnGlyGluTyr----------MetIleProCys------ 329

)y     1039 AATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCTGAGAGATGAGAAC 1098
Db      330 ---------GluLysValSerThrLeuProAlaIleThrLeuLysLeu---------- 342

Qy     1099 TCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCAGCCCATGATGGGG 1158
Db      343 GlyGlyLysGlyTyrLysLeuSer------ProGluAspTyrThrLeuLysValSerGln 360

Qy     1159 GCCGGCCTGAATTATGAATGTTACCGATTC---------GGCATTTCCCCATCCACAAAT 1209
Db      361 AlaGlyLysThrLeuCysLeuSerGlyPheMetGlyMetAspIleProProProSerGly 380

Qy     1210 GCGCTG---GTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAGAGCC 1266
Db      381 ProLeuTrpIleLeuGlyAspValPheIleGlyArgTyrTyrThrValPheAspArgAsp 400

Qy     1267 CAGAAGAGGGTGGGCTTCGCA 1287
Db      401 AsnAsnArgValGlyPheAla 407

RESULT 15
```

```
KHMSD
cathepsin D (EC 3.4.23.5) precursor - mouse
C;Species: Mus musculus (house mouse)
C;Date: 31-Dec-1991 #sequence_revision 31-Dec-1991 #text_change 18-Jun-1999
C;Accession: I48278; S14704; S12587
R;Hetman, M.; Perschl, A.; Saftig, P.; Von Figura, K.; Peters, C.
DNA Cell Biol. 13, 419-427, 1994
A;Title: Mouse cathepsin D gene: molecular organization, characterization of the prom
A;Reference number: I48278; MUID:94280622; PMID:8011168
A;Accession: I48278
A;Status: preliminary; translated from GB/EMBL/DDBJ
A;Molecule type: DNA
A;Residues: 1-410 <RES>
A;Cross-references: EMBL:X68378; NID:g50302; PIDN:CAA48453.1; PID:g817945
R;Diedrich, J.F.; Staskus, K.A.; Retzel, E.F.; Haase, A.T.
Nucleic Acids Res. 18, 7184, 1990
A;Title: Nucleotide sequence of a cDNA encoding mouse cathepsin D.
A;Reference number: S14704; MUID:91088345; PMID:2263503
A;Accession: S14704
A;Molecule type: mRNA
A;Residues: 1-410 <DIE>
A;Cross-references: EMBL:X53337; NID:g50300; PIDN:CAA37423.1; PID:g50301
R;Grusby, M.J.; Mitchell, S.C.; Glimcher, L.H.
Nucleic Acids Res. 18, 4008, 1990
A;Title: Molecular cloning of mouse cathepsin D.
A;Reference number: S12587; MUID:90326544; PMID:2374732
A;Accession: S12587
A;Molecule type: mRNA
A;Residues: 1-410 <GRU>
A;Cross-references: EMBL:X52886; NID:g50298; PIDN:CAA37067.1; PID:g50299
C;Genetics:
A;Introns: 23/2; 76/3; 118/1; 157/3; 233/2; 274/2; 322/3; 355/3
C;Function:
A;Description: limited specificity endopeptidase
A;Pathway: intracellular protein degradation
C;Superfamily: pepsin
C;Keywords: aspartic proteinase; glycoprotein; hydrolase; lysosome; protein degradati
F;1-20/Domain: signal sequence #status predicted <SIG>
F;21-64/Domain: propeptide #status predicted <PRO>
F;65-410/Product: cathepsin D, single-chain form #status predicted <MAT>
F;91-160,110-117,284-288,327-364/Disulfide bonds: #status predicted
F;97,293/Active site: Asp #status predicted
F;134,261/Binding site: carbohydrate (Asn) (covalent) #status predicted

Alignment Scores:
Pred. No.:                4.84e-14        Length:          410
Score:                    306.50          Matches:         103
Percent Similarity:       44.53%          Conservative:    64
Best Local Similarity:    27.47%          Mismatches:      123
Query Match:              9.18%           Indels:          85
DB:                       1               Gaps:            15

US-09-806-194-1 (1-1804) x KHMSD (1-410)

Qy      274 TACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCTACAGATTCTCGTTGACACT 333
Db       79 TyrTyrGlyAspIleGlyIleGlyThrProProGlnCysPheThrValValPheAspThr 98

Qy      334 GGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTACATAGACACG--------- 384
Db       99 GlySerSerAsnLeuTrpValProSerIleHisCysLysIleLeuAspIleAlaCysTrp 118

Qy      385 ---------TACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGTC 435
Db      119 ValHisHisLysTyrAsnSerAspLysSerSerThrTyrValLysAsnGlyThrSerPhe 138

Qy      436 ACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCATC 495
Db      139 AspIleHisTyrGlySerGlySerLeuSerGlyTyrLeuSerGlnAspThrValSerVal 158

Qy      496 CCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT---------ATTTTTGAA 546
Db      159 ProCysLysSerAspGlnSerLysAlaArgGlyIleLysValGluLysGlnIlePhe--- 177
```



Search completed: September 25, 2003, 11:09:36
Job time : 91 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 11:07:24 ; Search time 106.5 Seconds
(without alignments)
5125.972 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggcccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 566894 seqs, 151307093 residues

Total number of hits satisfying chosen parameters: 1133788

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095458_11190/app_query.fasta_1.1991
-DB=Published_Applications_AA -QFMT=fastan -SUFFIX=rapb -MINMATCH=0.1
-LOOPCL=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62
-TRANS=human40.cdi -LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100
-THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0
-MAXLEN=2000000000 -USER=US09806194_@CGN_1_1_118_@runat_25092003_095458_11190
-NCPU=6 -ICPU=3 -NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100
-LONGLOG -DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5
-FGAPOP=6 -FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Published_Applications_AA:*
1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep:*
7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep:*
10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep:*
11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep:*
12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep:*
14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep:*
15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep:*
16: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	2687	80.5	518	9	US-09-794-927-2	Sequence 2, Appli
2	2687	80.5	518	9	US-09-795-847-2	Sequence 2, Appli
3	2687	80.5	518	9	US-09-794-743-2	Sequence 2, Appli
4	2687	80.5	518	9	US-09-794-748-2	Sequence 2, Appli
5	2687	80.5	518	9	US-09-794-925-2	Sequence 2, Appli
6	2687	80.5	518	9	US-09-215-450-19	Sequence 19, Appl
7	2687	80.5	518	9	US-09-681-442-2	Sequence 2, Appli
8	2687	80.5	518	10	US-09-978-295A-196	Sequence 196, App
9	2687	80.5	518	10	US-09-886-143-2	Sequence 2, Appli
10	2687	80.5	518	10	US-09-978-697-196	Sequence 196, App
11	2687	80.5	518	10	US-09-978-192A-196	Sequence 196, App
12	2687	80.5	518	10	US-09-999-832A-196	Sequence 196, App
13	2687	80.5	518	11	US-09-978-189-196	Sequence 196, App
14	2687	80.5	518	11	US-09-978-608A-196	Sequence 196, App
15	2687	80.5	518	11	US-09-978-585A-196	Sequence 196, App
16	2687	80.5	518	11	US-09-978-191A-196	Sequence 196, App
17	2687	80.5	518	11	US-09-978-403A-196	Sequence 196, App
18	2687	80.5	518	11	US-09-978-564A-196	Sequence 196, App
19	2687	80.5	518	11	US-09-999-833A-196	Sequence 196, App
20	2687	80.5	518	11	US-09-981-915A-196	Sequence 196, App
21	2687	80.5	518	11	US-09-978-824-196	Sequence 196, App
22	2687	80.5	518	11	US-09-918-585A-196	Sequence 196, App
23	2687	80.5	518	11	US-09-978-423A-196	Sequence 196, App
24	2687	80.5	518	11	US-09-978-193A-196	Sequence 196, App
25	2687	80.5	518	11	US-09-669-414-2	Sequence 2, Appli
26	2687	80.5	518	11	US-09-999-830A-196	Sequence 196, App
27	2687	80.5	518	11	US-09-978-757A-196	Sequence 196, App
28	2687	80.5	518	11	US-09-978-187B-196	Sequence 196, App
29	2687	80.5	518	11	US-09-548-366-2	Sequence 2, Appli
30	2687	80.5	518	11	US-09-978-643A-196	Sequence 196, App
31	2687	80.5	518	12	US-09-978-375A-196	Sequence 196, App
32	2687	80.5	518	12	US-09-978-188A-196	Sequence 196, App
33	2687	80.5	518	12	US-09-978-298A-196	Sequence 196, App
34	2687	80.5	518	12	US-10-143-031A-196	Sequence 196, App
35	2687	80.5	518	12	US-10-002-967A-196	Sequence 196, App
36	2687	80.5	518	12	US-10-017-083A-196	Sequence 196, App
37	2687	80.5	518	12	US-10-143-030A-196	Sequence 196, App
38	2687	80.5	518	12	US-10-199-672-72	Sequence 72, Appl
39	2687	80.5	518	12	US-10-187-749-72	Sequence 72, Appl
40	2687	80.5	518	12	US-10-194-457-72	Sequence 72, Appl
41	2687	80.5	518	12	US-10-145-128A-196	Sequence 196, App
42	2687	80.5	518	12	US-10-184-642-72	Sequence 72, Appl
43	2687	80.5	518	12	US-10-196-747-72	Sequence 72, Appl
44	2687	80.5	518	12	US-10-173-689-72	Sequence 72, Appl
45	2687	80.5	518	12	US-10-173-690-72	Sequence 72, Appl

ALIGNMENTS

RESULT 1
US-09-794-927-2
; Sequence 2, Application US/09794927
; Patent No. US20010016324A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Riqiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
; TITLE OF INVENTION: USES
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280PG
; CURRENT APPLICATION NUMBER: US/09/794,927
; CURRENT FILING DATE: 2001-02-27
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23

```
   PRIOR APPLICATION NUMBER:  60/101,594
   PRIOR FILING DATE:  1998-09-24
   NUMBER OF SEQ ID NOS:  73
   SOFTWARE:  PatentIn Ver. 2.0
  SEQ ID NO 2
    LENGTH: 518
    TYPE: PRT
    ORGANISM: Homo sapiens
US-09-794-927-2

Alignment Scores:
Pred. No.:              1.34e-188       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     9               Gaps:           0

US-09-806-194-1 (1-1804) x US-09-794-927-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACICTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTCAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 2
US-09-795-847-2
; Sequence 2, Application US/09795847
; Patent No. US20010018208A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280DE
;  CURRENT APPLICATION NUMBER: US/09/795,847
;  CURRENT FILING DATE:  2001-02-28
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
```

```
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-795-847-2

Alignment Scores:
Pred. No.:                1.34e-188       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       9               Gaps:            0

US-09-806-194-1 (1-1804) x US-09-795-847-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCGC 60
              ||||||||||||||||||||||||||||||||||||||||| |||||:|||| ||| .
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              |||||||||||||||||||||||||||||||||||||||||||||| |||||||||:|||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCCTGGACGGGCTTCGTTGGGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGGCTAGCT 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              |||||||||||||||.|||||||||||||||||||||||||||||||||||||.|||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 3
US-09-794-743-2
; Sequence 2, Application US/09794743
; Patent No. US20010021391A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280BC
;  CURRENT APPLICATION NUMBER: US/09/794,743
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
```

```
  PRIOR APPLICATION NUMBER:  09/404,133
  PRIOR FILING DATE:  1999-09-23
  PRIOR APPLICATION NUMBER:  PCT/US99/20881
  PRIOR FILING DATE:  1999-09-23
  PRIOR APPLICATION NUMBER:  60/101,594
  PRIOR FILING DATE:  1998-09-24
  NUMBER OF SEQ ID NOS:  73
  SOFTWARE:  PatentIn Ver. 2.0
 SEQ ID NO 2
   LENGTH: 518
   TYPE: PRT
   ORGANISM: Homo sapiens
US-09-794-743-2

Alignment Scores:
Pred. No.:              1.34e-188     Length:          518
Score:                  2687.00       Matches:         518
Percent Similarity:     100.00%       Conservative:    0
Best Local Similarity:  100.00%       Mismatches:      0
Query Match:            80.47%        Indels:          0
DB:                     9             Gaps:            0

US-09-806-194-1 (1-1804) x US-09-794-743-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla  20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn  40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu  60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGGCATC 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet  80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 4
US-09-794-748-2
; Sequence 2, Application US/09794748
; Patent No. US20020037315A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND
;  TITLE OF INVENTION:  USES
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280JL
;  CURRENT APPLICATION NUMBER: US/09/794,748
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
;  PRIOR FILING DATE:  1999-10-13
```

```
;  PRIOR APPLICATION NUMBER:  60/155,493
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  09/404,133
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  PCT/US99/20881
;  PRIOR FILING DATE:  1999-09-23
;  PRIOR APPLICATION NUMBER:  60/101,594
;  PRIOR FILING DATE:  1998-09-24
;  NUMBER OF SEQ ID NOS:  73
;  SOFTWARE:  PatentIn Ver. 2.0
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-794-748-2

Alignment Scores:
Pred. No.:                1.34e-188       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       9               Gaps:            0

US-09-806-194-1 (1-1804) x US-09-794-748-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC  60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla  20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC  120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn  40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTC  180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu  60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG  240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet  80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC  300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA  360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC  420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA  480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT  540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT  600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA  660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT  720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA  780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA  840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC  900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG  960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 5
US-09-794-925-2
; Sequence 2, Application US/09794925
; Patent No. US20020064819A1
; GENERAL INFORMATION:
;  APPLICANT:  Gurney, Mark E.
;  APPLICANT:  Bienkowski, Michael J.
;  APPLICANT:  Heinrikson, Robert L.
;  APPLICANT:  Parodi, Luis A.
;  APPLICANT:  Yan, Riqiang
;  TITLE OF INVENTION:  ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND U
;  TITLE OF INVENTION:  THEREFOR
;  FILE REFERENCE:  28341/6280HI
;  CURRENT APPLICATION NUMBER: US/09/794,925
;  CURRENT FILING DATE:  2001-02-27
;  PRIOR APPLICATION NUMBER:  09/416,901
```

```
  PRIOR FILING DATE:  1999-10-13
  PRIOR APPLICATION NUMBER:  60/155,493
  PRIOR FILING DATE:  1999-09-23
  PRIOR APPLICATION NUMBER:  09/404,133
  PRIOR FILING DATE:  1999-09-23
  PRIOR APPLICATION NUMBER:  PCT/US99/20881
  PRIOR FILING DATE:  1999-09-23
  PRIOR APPLICATION NUMBER:  60/101,594
  PRIOR FILING DATE:  1998-09-24
  NUMBER OF SEQ ID NOS:  73
  SOFTWARE:  PatentIn Ver. 2.0
SEQ ID NO 2
  LENGTH: 518
  TYPE: PRT
  ORGANISM: Homo sapiens
US-09-794-925-2

Alignment Scores:
red. No.:              1.34e-188     Length:        518
core:                  2687.00       Matches:       518
ercent Similarity:     100.00%       Conservative:  0
est Local Similarity:  100.00%       Mismatches:    0
uery Match:            80.47%        Indels:        0
B:                     9             Gaps:          0

S-09-806-194-1 (1-1804) x US-09-794-925-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTCCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 6
US-09-215-450-19
; Sequence 19, Application US/09215450
; Patent No. US20020068278A1
; GENERAL INFORMATION:
;  APPLICANT: Giese, Klaus
;  APPLICANT: Xin, Hong
;  TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER REGULATED GENES
;  FILE REFERENCE: 1451.100 / 210030.447
;  CURRENT APPLICATION NUMBER: US/09/215,450
;  CURRENT FILING DATE: 1998-12-17
;  NUMBER OF SEQ ID NOS: 27
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 19
;   LENGTH: 518
```

TYPE: PRT
ORGANISM: human
09-215-450-19

gnment Scores:

ed. No.:	1.34e-188	Length:	518
ore:	2687.00	Matches:	518
rcent Similarity:	100.00%	Conservative:	0
st Local Similarity:	100.00%	Mismatches:	0
ery Match:	80.47%	Indels:	0
	9	Gaps:	0

-09-806-194-1 (1-1804) x US-09-215-450-19 (1-518)

```
          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

        121 CGCGTAGTTGCGCCCACCCCGGGATCCGGGACCCCTGCCGAGGGCCACGCCGACGGCTTG 180
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGCCCATG 240
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPheAspSerLeuValThr 220

        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACCCAGACAAGGCC 900
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProProArgCysGlnArgArgPro 500

       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 7
US-09-681-442-2
; Sequence 2, Application US/09681442
; Patent No. US20020081634A1
; GENERAL INFORMATION:
; APPLICANT: Gurney, Mark E.
; APPLICANT: Bienkowski, Michael J.
; APPLICANT: Heinrikson, Robert L.
; APPLICANT: Parodi, Luis A.
; APPLICANT: Yan, Rigiang
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR, AND US
; TITLE OF INVENTION: THEREFOR
; FILE REFERENCE: 28341/6280FG
; CURRENT APPLICATION NUMBER: US/09/681,442
; CURRENT FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 09/416,901
; PRIOR FILING DATE: 1999-10-13
; PRIOR APPLICATION NUMBER: 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2

LENGTH: 518
TYPE: PRT
ORGANISM: Homo sapiens
09-681-442-2

gnment Scores:

d. No.:	1.34e-188	Length:	518
ore:	2687.00	Matches:	518
cent Similarity:	100.00%	Conservative:	0
t Local Similarity:	100.00%	Mismatches:	0
ery Match:	80.47%	Indels:	0
	9	Gaps:	0

09-806-194-1 (1-1804) x US-09-681-442-2 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACICCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTIGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 8
US-09-978-295A-196
; Sequence 196, Application US/09978295A
; Patent No. US20020156006A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Elien
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann

; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C11
; CURRENT APPLICATION NUMBER: US/09/978,295A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07

PRIOR APPLICATION NUMBER: 60/084598
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084600
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084627
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/084643
PRIOR FILING DATE: 1998-05-07
PRIOR APPLICATION NUMBER: 60/085339
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085338
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085323
PRIOR FILING DATE: 1998-05-13
PRIOR APPLICATION NUMBER: 60/085582
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085700
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085689
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085579
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085580
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085573
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085704
PRIOR FILING DATE: 1998-05-15
PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	10	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-978-295A-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 9
US-09-886-143-2

```
; Sequence 2, Application US/09886143
; Patent No. US20020159991A1
; GENERAL INFORMATION:
;  APPLICANT: Cordell, Barbara
;  APPLICANT:  Schimmoller, Frauke
;  APPLICANT:  Liu, Yu-Wang
;  APPLICANT:  Quon, Diana Hom
;  TITLE OF INVENTION: Modulation of A  Levels by
;  TITLE OF INVENTION:  Secretase BACE2
;  FILE REFERENCE: SCIOS.022A
;  CURRENT APPLICATION NUMBER: US/09/886,143
;  CURRENT FILING DATE:  2001-06-20
;  PRIOR APPLICATION NUMBER: 60/215,729
;  PRIOR FILING DATE: 2000-06-28
;  NUMBER OF SEQ ID NOS: 6
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-886-143-2

Alignment Scores:
Pred. No.:                1.34e-188      Length:          518
Score:                    2687.00        Matches:         518
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              80.47%         Indels:          0
DB:                       10             Gaps:            0

US-09-806-194-1 (1-1804) x US-09-886-143-2 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 10
US-09-978-697-196
; Sequence 196, Application US/09978697
; Patent No. US20020169284A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
```

; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C27
; CURRENT APPLICATION NUMBER: US/09/978,697
; CURRENT FILING DATE: 2001-10-16
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/C79656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559

; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	10	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-978-697-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
```

```
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              |||||||||||||||||||||||||||||||||||||||||||||||| |||||||||||
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 11
US-09-978-192A-196
; Sequence 196, Application US/09978192A
; Patent No. US20020177553A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C9
; CURRENT APPLICATION NUMBER: US/09/978,192A
; CURRENT FILING DATE: 2001-10-15
; PRIOR APPLICATION NUMBER: 09/918585
; PRIOR FILING DATE: 2001-07-30
; PRIOR APPLICATION NUMBER: 60/062250
; PRIOR FILING DATE: 1997-10-17
; PRIOR APPLICATION NUMBER: 60/064249
; PRIOR FILING DATE: 1997-11-03
; PRIOR APPLICATION NUMBER: 60/065311
; PRIOR FILING DATE: 1997-11-13
; PRIOR APPLICATION NUMBER: 60/066364
; PRIOR FILING DATE: 1997-11-21
; PRIOR APPLICATION NUMBER: 60/077450
; PRIOR FILING DATE: 1998-03-10
; PRIOR APPLICATION NUMBER: 60/077632
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077641
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077649
; PRIOR FILING DATE: 1998-03-11
; PRIOR APPLICATION NUMBER: 60/077791
; PRIOR FILING DATE: 1998-03-12
; PRIOR APPLICATION NUMBER: 60/078004
; PRIOR FILING DATE: 1998-03-13
; PRIOR APPLICATION NUMBER: 60/078886
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078936
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078910
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/078939
; PRIOR FILING DATE: 1998-03-20
; PRIOR APPLICATION NUMBER: 60/079294
; PRIOR FILING DATE: 1998-03-25
; PRIOR APPLICATION NUMBER: 60/079656
; PRIOR FILING DATE: 1998-03-26
; PRIOR APPLICATION NUMBER: 60/079664
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079689
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079663
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079728
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079786
; PRIOR FILING DATE: 1998-03-27
; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22

; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084527
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	10	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-978-192A-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320
```

```
Qy       961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy      1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy      1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy      1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy      1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy      1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy      1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
            ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 12
US-09-999-832A-196
; Sequence 196, Application US/09999832A
; Publication No. US20020192706A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C63
;  CURRENT APPLICATION NUMBER: US/09/999,832A
;  CURRENT FILING DATE:  2001-10-24
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079920
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/079923
;  PRIOR FILING DATE: 1998-03-30
;  PRIOR APPLICATION NUMBER: 60/080105
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080107
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080165
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080194
;  PRIOR FILING DATE: 1998-03-31
;  PRIOR APPLICATION NUMBER: 60/080327
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080328
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080333
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/080334
;  PRIOR FILING DATE: 1998-04-01
;  PRIOR APPLICATION NUMBER: 60/081070
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081049
;  PRIOR FILING DATE: 1998-04-08
;  PRIOR APPLICATION NUMBER: 60/081071
```

; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	10	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-999-832A-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200
```

```
Qy       601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db       201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy       661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db       221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy       721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db       241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy       781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db       261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy       841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGACAGTATAACGCAGACAAGGCC 900
Db       281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy       901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db       301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy       961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db       321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy      1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db       341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy      1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db       361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy      1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db       381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy      1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db       401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy      1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db       421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy      1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTGCCCGCT 1380
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTGTGTGGA 1440
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCGC 1500
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 13
US-09-978-189-196
; Sequence 196, Application US/09978189
; Publication No. US20030004102A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C7
;  CURRENT APPLICATION NUMBER: US/09/978,189
;  CURRENT FILING DATE:  2001-10-15
;  PRIOR APPLICATION NUMBER: 09/918585
;  PRIOR FILING DATE: 2001-07-30
;  PRIOR APPLICATION NUMBER: 60/062250
;  PRIOR FILING DATE: 1997-10-17
;  PRIOR APPLICATION NUMBER: 60/064249
;  PRIOR FILING DATE: 1997-11-03
;  PRIOR APPLICATION NUMBER: 60/065311
;  PRIOR FILING DATE: 1997-11-13
;  PRIOR APPLICATION NUMBER: 60/066364
;  PRIOR FILING DATE: 1997-11-21
;  PRIOR APPLICATION NUMBER: 60/077450
;  PRIOR FILING DATE: 1998-03-10
;  PRIOR APPLICATION NUMBER: 60/077632
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077641
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077649
;  PRIOR FILING DATE: 1998-03-11
;  PRIOR APPLICATION NUMBER: 60/077791
;  PRIOR FILING DATE: 1998-03-12
;  PRIOR APPLICATION NUMBER: 60/078004
;  PRIOR FILING DATE: 1998-03-13
;  PRIOR APPLICATION NUMBER: 60/078886
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078936
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078910
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/078939
;  PRIOR FILING DATE: 1998-03-20
;  PRIOR APPLICATION NUMBER: 60/079294
;  PRIOR FILING DATE: 1998-03-25
;  PRIOR APPLICATION NUMBER: 60/079656
;  PRIOR FILING DATE: 1998-03-26
;  PRIOR APPLICATION NUMBER: 60/079664
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079689
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079663
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079728
;  PRIOR FILING DATE: 1998-03-27
;  PRIOR APPLICATION NUMBER: 60/079786
;  PRIOR FILING DATE: 1998-03-27
```

; PRIOR APPLICATION NUMBER: 60/079920
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/079923
; PRIOR FILING DATE: 1998-03-30
; PRIOR APPLICATION NUMBER: 60/080105
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080107
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080165
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080194
; PRIOR FILING DATE: 1998-03-31
; PRIOR APPLICATION NUMBER: 60/080327
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080328
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080333
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/080334
; PRIOR FILING DATE: 1998-04-01
; PRIOR APPLICATION NUMBER: 60/081070
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081049
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081071
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081195
; PRIOR FILING DATE: 1998-04-08
; PRIOR APPLICATION NUMBER: 60/081203
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081229
; PRIOR FILING DATE: 1998-04-09
; PRIOR APPLICATION NUMBER: 60/081955
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081817
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081819
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081952
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/081838
; PRIOR FILING DATE: 1998-04-15
; PRIOR APPLICATION NUMBER: 60/082568
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082569
; PRIOR FILING DATE: 1998-04-21
; PRIOR APPLICATION NUMBER: 60/082704
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082804
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082700
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082797
; PRIOR FILING DATE: 1998-04-22
; PRIOR APPLICATION NUMBER: 60/082796
; PRIOR FILING DATE: 1998-04-23
; PRIOR APPLICATION NUMBER: 60/083336
; PRIOR FILING DATE: 1998-04-27
; PRIOR APPLICATION NUMBER: 60/083322
; PRIOR FILING DATE: 1998-04-28
; PRIOR APPLICATION NUMBER: 60/083392
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083495
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083496
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083499
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083545
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083554
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083558
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083559
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083500
; PRIOR FILING DATE: 1998-04-29
; PRIOR APPLICATION NUMBER: 60/083742
; PRIOR FILING DATE: 1998-04-30
; PRIOR APPLICATION NUMBER: 60/084366
; PRIOR FILING DATE: 1998-05-05
; PRIOR APPLICATION NUMBER: 60/084414
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084441
; PRIOR FILING DATE: 1998-05-06
; PRIOR APPLICATION NUMBER: 60/084637
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084639
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084640
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084598
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084600
; PRIOR FILING DATE: 1998-5-07
; PRIOR APPLICATION NUMBER: 60/084627
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/084643
; PRIOR FILING DATE: 1998-05-07
; PRIOR APPLICATION NUMBER: 60/085339
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085338
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085323
; PRIOR FILING DATE: 1998-05-13
; PRIOR APPLICATION NUMBER: 60/085582
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085700
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085689
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085579
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085580
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085573
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085704
; PRIOR FILING DATE: 1998-05-15
; PRIOR APPLICATION NUMBER: 60/085697

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	11	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-978-189-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy       241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy       301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGGCGTGGCAGGA 360
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy       361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy       421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCCTCGTTGGGGAA 480
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy       481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy       541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy       601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy       661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy       721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy       781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy       841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy       901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy       961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy      1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProLysIleSerIle 360

Qy      1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy      1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy      1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy      1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy      1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAAA 1554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 14
US-09-978-608A-196
; Sequence 196, Application US/09978608A
; Publication No. US20030045462A1
; GENERAL INFORMATION:
;  APPLICANT: Ashkenazi, Avi
;  APPLICANT:  Baker Kevin P.
;  APPLICANT:  Botstein, David
;  APPLICANT:  Desnoyers, Luc
;  APPLICANT:  Eaton, Dan
;  APPLICANT:  Ferrara, Napoleon
;  APPLICANT:  Filvaroff, Ellen
;  APPLICANT:  Fong, Sherman
;  APPLICANT:  Gao, Wei-Qiang
;  APPLICANT:  Gerber, Hanspeter
;  APPLICANT:  Gerritsen, Mary E.
;  APPLICANT:  Goddard, Audrey
;  APPLICANT:  Godowski, Paul J.
;  APPLICANT:  Grimaldi, J. Christopher
;  APPLICANT:  Gurney, Austin L.
;  APPLICANT:  Hillan, Kenneth J
;  APPLICANT:  Kljavin, Ivar J.
;  APPLICANT:  Kuo, Sophia S.
;  APPLICANT:  Napier, Mary A.
;  APPLICANT:  Pan, James;
;  APPLICANT:  Paoni, Nicholas F.
;  APPLICANT:  Roy, Margaret Ann
;  APPLICANT:  Shelton, David L.
;  APPLICANT:  Stewart, Timothy A.
;  APPLICANT:  Tumas, Daniel
;  APPLICANT:  Williams, P. Mickey
;  APPLICANT:  Wood, William I.
;  TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
;  TITLE OF INVENTION:  Acids Encoding the Same
;  FILE REFERENCE: P2630P1C22
;  CURRENT APPLICATION NUMBER: US/09/978,608A
;  CURRENT FILING DATE:  2001-10-16
;  NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 196
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo sapien
US-09-978-608A-196

Alignment Scores:
Pred. No.:              1.34e-188       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     11              Gaps:           0

US-09-806-194-1 (1-1804) x US-09-978-608A-196 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
```

```
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCCACGCCGACGGCTTG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCAICAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGAICTGGGACCAACGGAGGTAGTCITGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 15
US-09-978-585A-196
; Sequence 196, Application US/09978585A
; Publication No. US20030049633A1
; GENERAL INFORMATION:
; APPLICANT: Ashkenazi, Avi
; APPLICANT: Baker Kevin P.
; APPLICANT: Botstein, David
; APPLICANT: Desnoyers, Luc
; APPLICANT: Eaton, Dan
; APPLICANT: Ferrara, Napoleon
; APPLICANT: Filvaroff, Ellen
; APPLICANT: Fong, Sherman
; APPLICANT: Gao, Wei-Qiang
; APPLICANT: Gerber, Hanspeter
; APPLICANT: Gerritsen, Mary E.
; APPLICANT: Goddard, Audrey
; APPLICANT: Godowski, Paul J.
; APPLICANT: Grimaldi, J. Christopher
; APPLICANT: Gurney, Austin L.
; APPLICANT: Hillan, Kenneth J
; APPLICANT: Kljavin, Ivar J.
; APPLICANT: Kuo, Sophia S.
; APPLICANT: Napier, Mary A.
; APPLICANT: Pan, James;
; APPLICANT: Paoni, Nicholas F.
; APPLICANT: Roy, Margaret Ann
; APPLICANT: Shelton, David L.
; APPLICANT: Stewart, Timothy A.
; APPLICANT: Tumas, Daniel
; APPLICANT: Williams, P. Mickey
; APPLICANT: Wood, William I.
; TITLE OF INVENTION: Secreted and Transmembrane Polypeptides and Nucleic
; TITLE OF INVENTION: Acids Encoding the Same
; FILE REFERENCE: P2630P1C15
; CURRENT APPLICATION NUMBER: US/09/978,585A
; CURRENT FILING DATE: 2001-10-16
; NUMBER OF SEQ ID NOS: 624
; Prior Application removed - See File Wrapper or Palm
; SEQ ID NO 196
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapien

US-09-978-585A-196

Alignment Scores:

Pred. No.:	1.34e-188	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	11	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-978-585A-196 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

Search completed: September 25, 2003, 11:34:19
Job time : 139.5 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:55:54 ; Search time 33.5 Seconds
(without alignments)
4556.943 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 328717 seqs, 42310858 residues

Total number of hits satisfying chosen parameters: 657434

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095456_11048/app_query.fasta_1.1991
-DB=Issued_Patents_AA -QFMT=fastan -SUFFIX=rai -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09806194_@CGN_1_1_42_@runat_25092003_095456_11048 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued_Patents_AA:*
1: /cgn2_6/ptodata/1/iaa/5A_COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/5B_COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/6A_COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/6B_COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/PCTUS_COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2687	80.5	518	3	US-08-999-723-2	Sequence 2, Appli
2	2687	80.5	518	3	US-09-434-427-2	Sequence 2, Appli
3	2687	80.5	518	4	US-09-548-372D-2	Sequence 2, Appli
4	2687	80.5	518	4	US-09-548-367D-2	Sequence 2, Appli
5	2687	80.5	518	4	US-09-551-853D-2	Sequence 2, Appli
6	2395	71.7	514	3	US-09-717-432-2	Sequence 2, Appli
7	2395	71.7	514	4	US-09-912-484-2	Sequence 2, Appli
8	1186.5	35.5	501	4	US-09-713-158-2	Sequence 2, Appli
9	1185	35.5	501	4	US-09-548-372D-8	Sequence 8, Appli
10	1185	35.5	501	4	US-09-548-367D-8	Sequence 8, Appli
11	1185	35.5	501	4	US-09-551-853D-8	Sequence 8, Appli
12	1178.5	35.3	501	4	US-09-548-372D-4	Sequence 4, Appli
13	1178.5	35.3	501	4	US-09-548-367D-4	Sequence 4, Appli
14	1178.5	35.3	501	4	US-09-551-853D-4	Sequence 4, Appli
15	1175.5	35.2	503	4	US-09-604-608-3	Sequence 3, Appli
16	1175	35.2	488	4	US-09-604-608-2	Sequence 2, Appli
17	1172.5	35.1	501	4	US-09-009-191-2	Sequence 2, Appli
18	1160.5	34.8	774	4	US-09-009-191-4	Sequence 4, Appli
19	1139	34.1	453	4	US-09-548-372D-30	Sequence 30, Appl
20	1139	34.1	453	4	US-09-548-367D-30	Sequence 30, Appl
21	1139	34.1	453	4	US-09-551-853D-30	Sequence 30, Appl
22	1139	34.1	459	4	US-09-548-372D-32	Sequence 32, Appl
23	1139	34.1	459	4	US-09-548-367D-32	Sequence 32, Appl
24	1139	34.1	459	4	US-09-551-853D-32	Sequence 32, Appl
25	1130.5	33.9	446	4	US-09-548-372D-22	Sequence 22, Appl
26	1130.5	33.9	446	4	US-09-548-367D-22	Sequence 22, Appl
27	1130.5	33.9	446	4	US-09-551-853D-22	Sequence 22, Appl
28	1129	33.8	459	4	US-09-548-372D-24	Sequence 24, Appl
29	1129	33.8	459	4	US-09-548-367D-24	Sequence 24, Appl
30	1129	33.8	459	4	US-09-551-853D-24	Sequence 24, Appl
31	1127	33.8	433	4	US-09-548-372D-26	Sequence 26, Appl
32	1127	33.8	433	4	US-09-548-367D-26	Sequence 26, Appl
33	1127	33.8	433	4	US-09-551-853D-26	Sequence 26, Appl
34	1123	33.6	425	4	US-09-548-372D-28	Sequence 28, Appl
35	1123	33.6	425	4	US-09-548-367D-28	Sequence 28, Appl
36	1123	33.6	425	4	US-09-551-853D-28	Sequence 28, Appl
37	1071.5	32.1	476	4	US-09-548-372D-73	Sequence 73, Appl
38	1071.5	32.1	476	4	US-09-548-367D-73	Sequence 73, Appl
39	1071.5	32.1	476	4	US-09-551-853D-73	Sequence 73, Appl
40	1068	32.0	476	4	US-09-548-372D-6	Sequence 6, Appli
41	1068	32.0	476	4	US-09-548-367D-6	Sequence 6, Appli
42	1068	32.0	476	4	US-09-551-853D-6	Sequence 6, Appli
43	1028.5	30.8	428	4	US-09-548-372D-51	Sequence 51, Appl
44	1028.5	30.8	428	4	US-09-548-367D-51	Sequence 51, Appl
45	1028.5	30.8	428	4	US-09-551-853D-51	Sequence 51, Appl

ALIGNMENTS

RESULT 1
US-08-999-723-2
; Sequence 2, Application US/08999723A
; Patent No. 6025180
; GENERAL INFORMATION:
; APPLICANT: Powell, David J.
; APPLICANT: Southan, Christopher
; APPLICANT: Chapman, Conrad G.
; APPLICANT: Evans, Joanne R.
; TITLE OF INVENTION: ASP1
; FILE REFERENCE: GH70262
; CURRENT APPLICATION NUMBER: US/08/999,723A
; CURRENT FILING DATE: 1997-10-06
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-08-999-723-2

Alignment Scores:

Pred. No.:	9.62e-227	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	3	Gaps:	0

US-09-806-194-1 (1-1804) x US-08-999-723-2 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
```

```
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||| |||||| ||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              |||||||||||||||||||||||||||||||||||||||||||||| ||||||| ||| ||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGly 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 2
US-09-434-427-2
; Sequence 2, Application US/09434427
; Patent No. 6162630
; GENERAL INFORMATION:
;  APPLICANT: POWELL, DAVID J.
;  APPLICANT: SOUTHAN, CHRISTOPHER
;  APPLICANT: CHAPMAN, CONRAD G.
;  APPLICANT: EVANS, JOANNE R.
;  TITLE OF INVENTION: ASP1
;  FILE REFERENCE: GH-70262-D1
;  CURRENT APPLICATION NUMBER: US/09/434,427
;  CURRENT FILING DATE: 1999-11-04
;  EARLIER APPLICATION NUMBER: US 08/999,723
;  EARLIER FILING DATE: 1997-10-06
;  EARLIER APPLICATION NUMBER: UK 9626022.9
;  EARLIER FILING DATE: 1996-12-14
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: HOMO SAPIENS
US-09-434-427-2

Alignment Scores:
Pred. No.:                9.62e-227      Length:          518
Score:                    2687.00        Matches:         518
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              80.47%         Indels:          0
DB:                       3              Gaps:            0

US-09-806-194-1 (1-1804) x US-09-434-427-2 (1-518)

Qy           1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy          61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy         121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60
```

```
Qy       181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db        61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy       241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db        81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy       301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db       101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy       361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db       121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy       421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db       141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy       481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db       161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy       541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db       181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy       601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db       201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy       661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db       221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy       721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db       241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy       781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db       261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy       841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db       281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy       901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db       301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy       961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db       321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy      1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db       341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy      1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db       361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy      1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db       381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy      1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db       401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy      1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db       421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy      1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db       441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy      1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db       461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy      1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db       481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy      1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db       501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 3
US-09-548-372D-2
; Sequence 2, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND U
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-2

Alignment Scores:

Pred. No.:	9.62e-227	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	4	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-548-372D-2 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
```

```
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTCGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 4
US-09-548-367D-2
; Sequence 2, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND US
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280H
; CURRENT APPLICATION NUMBER: US/09/548,367D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 518
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-367D-2

Alignment Scores:

Pred. No.:	9.62e-227	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	4	Gaps:	0

US-09-806-194-1 (1-1804) x US-09-548-367D-2 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTC 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 5
US-09-551-853D-2
; Sequence 2, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND U
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280L
;  CURRENT APPLICATION NUMBER: US/09/551,853D
;  CURRENT FILING DATE:  2000-04-18
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: Patentln version 3.1
; SEQ ID NO 2
;   LENGTH: 518
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-551-853D-2

Alignment Scores:
Pred. No.:                9.62e-227       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       4               Gaps:            0

US-09-806-194-1 (1-1804) x US-09-551-853D-2 (1-518)

Qy           1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db           1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy          61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy         121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy         181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy         241 GTAGACAACCTGCAGGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy         301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy         361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140
```

```
Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGGCT 1380
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 6
US-09-717-432-2
; Sequence 2, Application US/09717432
; Patent No. 6291223
; GENERAL INFORMATION:
; APPLICANT: ZHU, YUAN
; APPLICANT: LI, XIAOTONG
; APPLICANT: CHRISTIE, GARY
; APPLICANT: POWELL, DAVID J.
; TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASP1)
; FILE REFERENCE: GP-70663
; CURRENT APPLICATION NUMBER: US/09/717,432
; CURRENT FILING DATE: 2000-11-21
; PRIOR APPLICATION NUMBER: 60/166,974
; PRIOR FILING DATE: 1999-11-23
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
; LENGTH: 514
; TYPE: PRT
; ORGANISM: MUS MUSCULUS
US-09-717-432-2

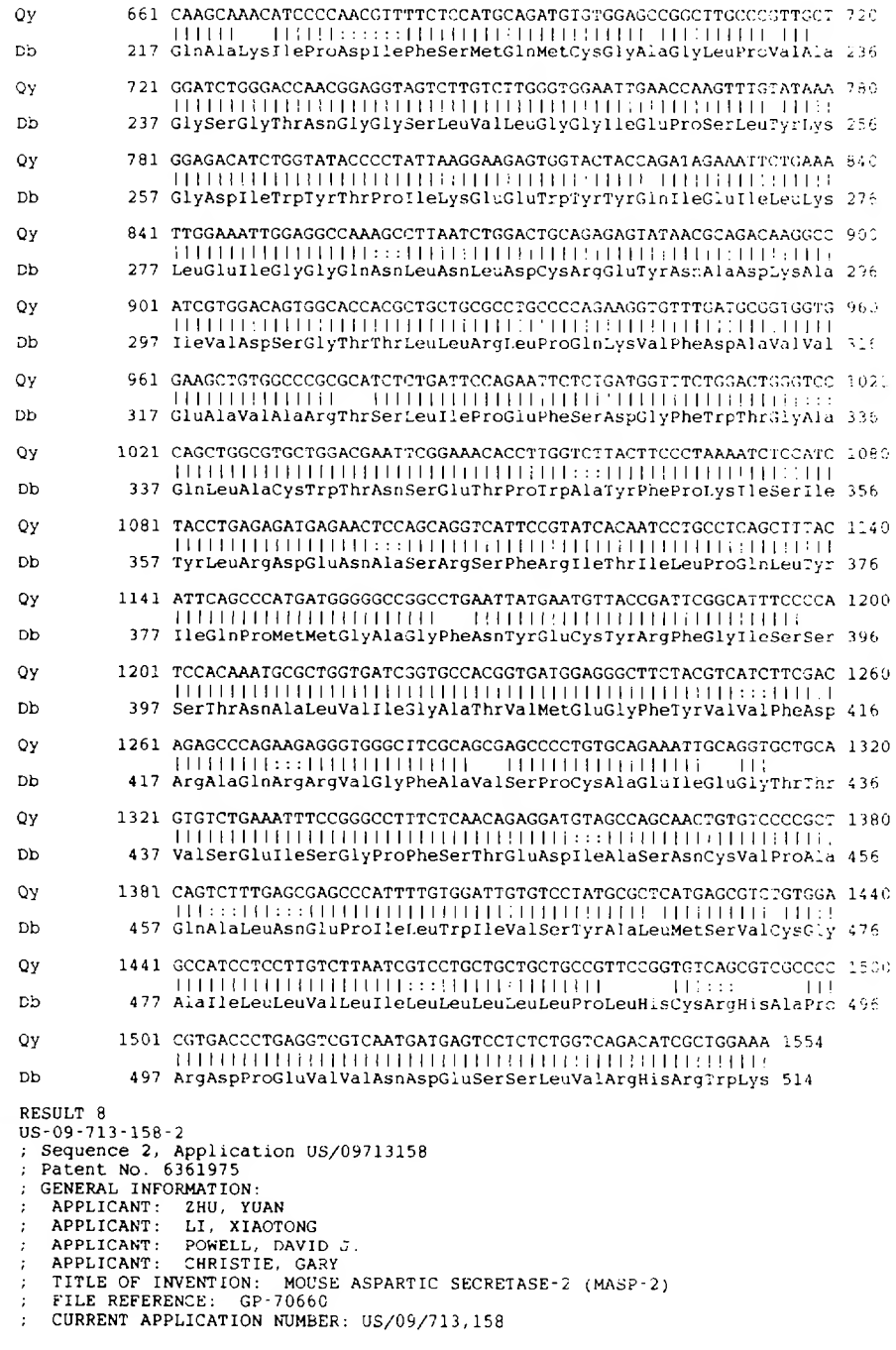
Alignment Scores:

Pred. No.:	3.43e-201	Length:	514
Score:	2395.00	Matches:	459
Percent Similarity:	92.47%	Conservative:	20
Best Local Similarity:	88.61%	Mismatches:	35
Query Match:	71.73%	Indels:	4
DB:	3	Gaps:	1

US-09-806-194-1 (1-1804) x US-09-717-432-2 (1-514)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuLeuValLeuAlaGlnTrpLeuLeuSerAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaGlyAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProValArgAlaThr----------AlaAsnPheLeuAlaMet 76

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 96

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db         97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176
```

```
Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||:|||||||||||||||||||||||||||||||||||||||
Db        177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||  ||||||||||.||||||||||||||||||||||||||||||.||||
Db        197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||  ||||||:::::|||||||||||||||||||||||||||||||||||||||||
Db        217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        257 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 276

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||:::||||||||||||||||||||||||||||||||||||||||
Db        277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316

Qy        961 GAAGCTGTGGCCCCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||  ||||||||||||||||||||||||||||||||||||||||||:::
Db        317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTCCCTAAAATCTCCATC 1080
              |||||||||||||||||||||||||||||||||||||:::||||.|||||||||| ||||:
Db        337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIle 356

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||:::||||||||||||||||||| ||||||||||||||||||||
Db        357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 376

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||.|||||||||||||||||||||||||||:|||| ||||||:|||
Db        397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||:|||||||||||||||  |||||||||||||||||  |||
Db        417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||:::||||.|||||||||||||||||
Db        437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              |||:::||||:::|||||||||||||||||||||||||||||||||||||||||||||||
Db        457 GlnAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 476

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||:::|||||||||||||||  ||||:::  |||
Db        477 AlaIleLeuLeuValLeuIleLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              |||||||||||||||||||||||||||||||||||||||||||||||||:|||||
Db        497 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 514

RESULT 7
US-09-912-484-2
; Sequence 2, Application US/09912484
; Patent No. 6358725
; GENERAL INFORMATION:
;  APPLICANT: Christie, Gary
;  APPLICANT:  Li, Xiaotong
;  APPLICANT:  Powell,  David J.
;  APPLICANT:  Zhu, Yuan
;  TITLE OF INVENTION: Mouse Aspartic Secretase-1 (mASP1)
;  FILE REFERENCE: GP-70663-D1
;  CURRENT APPLICATION NUMBER: US/09/912,484
;  CURRENT FILING DATE:  2001-07-25
;  PRIOR APPLICATION NUMBER: 60/166,974
;  PRIOR FILING DATE: 1999-11-23
;  PRIOR APPLICATION NUMBER: 09/717,432
;  PRIOR FILING DATE: 2000-11-21
;  NUMBER OF SEQ ID NOS: 2
;  SOFTWARE: FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;   LENGTH: 514
;   TYPE: PRT
;   ORGANISM: MUS MUSCULUS
US-09-912-484-2

Alignment Scores:
Pred. No.:                3.43e-201      Length:          514
Score:                    2395.00        Matches:         459
Percent Similarity:       92.47%         Conservative:    20
Best Local Similarity:    88.61%         Mismatches:      35
Query Match:              71.73%         Indels:          4
DB:                       4              Gaps:            1

US-09-806-194-1 (1-1804) x US-09-912-484-2 (1-514)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              |||||||||||   ||||||||||||||   :::||||||||||||||||||   |||
Db          1 MetGlyAlaLeuLeuArgAlaLeuLeuLeuLeuValLeuAlaGlnTrpLeuLeuSerAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
                 |||   ||||||||||||||||||||||||||||:::|||||||   ||||||||
Db         21 ValProAlaLeuAlaProAlaProPheThrLeuProLeuGlnValAlaGlyAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
                     :::      ||||||   |||||||||            ||||||||||||
Db         41 HisArgAlaSerAlaValProGlyLeuGlyThrProGluLeuProArgAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||      ||||:::            ||||||||||||||||
Db         61 AlaLeuAlaLeuGluProValArgAlaThr-----------AlaAsnPheLeuAlaMet 76

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||:||||||||||||||||||||||||||||||||||||||||||||||||||
Db         77 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 96

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||:::||||||||||||||||||||||||||||||||||||||||||||||
Db         97 ProProGlnLysValGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 116

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||:::|||  ||||||||||||  |||
Db        117 AlaProHisSerTyrIleAspThrTyrPheAspSerGluSerSerSerThrTyrHisSer 136

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        137 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 156

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||:::||||||||||||||||||||||||||||
Db        157 AspLeuValThrIleProLysGlyPheAsnSerSerPheLeuValAsnIleAlaThrIle 176

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        177 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 196

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||  ||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        197 TyrAlaAlaLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValAla 216
```



```
Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        217 GlnAlaLysIleProAspIlePheSerMetGlnMetCysGlyAlaGlyLeuProValAla 236

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        237 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 256

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        257 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 276

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        277 LeuGluIleGlyGlyGlnAsnLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 296

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        297 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 316

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        317 GluAlaValAlaArgThrSerLeuIleProGluPheSerAspGlyPheTrpThrGlyAla 336

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        337 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpAlaTyrPheProLysIleSerIle 356

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
Db        357 TyrLeuArgAspGluAsnAlaSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 376

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        377 IleGlnProMetMetGlyAlaGlyPheAsnTyrGluCysTyrArgPheGlyIleSerSer 396

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        397 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValValPheAsp 416

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        417 ArgAlaGlnArgArgValGlyPheAlaValSerProCysAlaGluIleGluGlyThrThr 436

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        437 ValSerGluIleSerGlyProPheSerThrGluAspIleAlaSerAsnCysValProAla 456

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        457 GlnAlaLeuAsnGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 476

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        477 AlaIleLeuLeuValLeuIleLeuLeuLeuLeuProLeuHisCysArgHisAlaPro 496

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        497 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 514

RESULT 8
US-09-713-158-2
; Sequence 2, Application US/09713158
; Patent No. 6361975
; GENERAL INFORMATION:
;  APPLICANT:  ZHU, YUAN
;  APPLICANT:  LI, XIAOTONG
;  APPLICANT:  POWELL, DAVID J.
;  APPLICANT:  CHRISTIE, GARY
;  TITLE OF INVENTION:  MOUSE ASPARTIC SECRETASE-2 (MASP-2)
;  FILE REFERENCE:  GP-70660
;  CURRENT APPLICATION NUMBER: US/09/713,158
;  CURRENT FILING DATE:  2000-11-15
;  PRIOR APPLICATION NUMBER:  60/165,800
;  PRIOR FILING DATE:  1999-11-16
;  NUMBER OF SEQ ID NOS:  2
;  SOFTWARE:  FastSEQ for Windows Version 3.0
; SEQ ID NO 2
;    LENGTH: 501
;    TYPE: PRT
;    ORGANISM: MUS MUSCULUS
US-09-713-158-2

Alignment Scores:
Pred. No.:                1.96e-95        Length:          501
Score:                    1186.50         Matches:         238
Percent Similarity:       62.09%          Conservative:    83
Best Local Similarity:    46.03%          Mismatches:      167
Query Match:              35.53%          Indels:          29
DB:                       4               Gaps:            8

US-09-806-194-1 (1-1804) x US-09-713-158-2 (1-501)

Qy         19 GCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC 78
Db          2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpVal---------GlySerGlyMetLeuPro 18

Qy         79 GCGCCCTTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC 123
Db         19 AlaGlnGlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla--------- 35

Qy        124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183
Db         36 ------GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------ 51

Qy        184 CTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCITGGCCATGGTA 243
Db         52 --------------------GluGluProGlyArgArgGlySerPheValGluMetVal 64

Qy        244 GACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCC 303
Db         65 AspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerPro 84

Qy        304 CCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACC 363
Db         85 ProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAla 104

Qy        364 CCGCACTCCTACATAGACACGTACTITGACACAGAGAGGTCTAGCACATACCGCTCCAAG 423
Db        105 ProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeu 124

Qy        424 GGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGAC 483
Db        125 ArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAsp 144

Qy        484 CTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTT 543
Db        145 LeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThr 164

Qy        544 GAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTAT 603
Db        165 GluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyr 184

Qy        604 GCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAA 663
Db        185 AlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGln 204

Qy        664 GCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT------ 717
Db        205 ThrHisIleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGln 224

Qy        718 ---GCTGGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTG 774
Db        225 ThrGluAlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeu 244
```

```
Qy        775 TATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATT 834
Db        245 TyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIle 264

Qy        835 CTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGAC 894
Db        265 ValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAsp 284

Qy        895 AAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCG 954
Db        285 LysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAla 304

Qy        955 GTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACT 1014
Db        305 AlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeu 324

Qy       1015 GGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATC 1074
Db        325 GlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIle 344

Qy       1075 TCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAG 1134
Db        345 SerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGln 364

Qy       1135 CTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGC 1191
Db        365 GlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAla 384

Qy       1192 ATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTC 1251
Db        385 ValSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrVal 404

Qy       1252 ATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCA 1311
Db        405 ValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAsp 424

Qy       1312 GGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGT 1371
Db        425 GluPheArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGly 444

Qy       1372 GTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGC 1431
Db        445 TyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAla 464

Qy       1432 GTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAG 1491
Db        465 IleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeu 483

Qy       1492 CGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        484 ArgCysLeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 503

RESULT 9
US-09-548-372D-8
; Sequence 8, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280I
;  CURRENT APPLICATION NUMBER: US/09/548,372D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1

; SEQ ID NO 8
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-548-372D-8

Alignment Scores:
Pred. No.:                2.65e-95        Length:          501
Score:                    1185.00         Matches:         237
Percent Similarity:       62.14%          Conservative:    83
Best Local Similarity:    46.02%          Mismatches:      169
Query Match:              35.49%          Indels:          26
DB:                       4               Gaps:            7

US-09-806-194-1 (1-1804) x US-09-548-372D-8 (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla--------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
```

```
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960

Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020

Db        307 LysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGlu 326

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080

Db        327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePheProValIleSerLeu 346

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140

Db        347 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 366

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197

Db        367 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257

Db        387 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317

Db        407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377

Db        427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437

Db        447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT--- 1494

Db        467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536

Db        486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 10
US-09-548-367D-8
; Sequence 8, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: Mus musculus
US-09-548-367D-8
```

```
Alignment Scores:
Pred. No.:                2.65e-95      Length:          501
Score:                    1185.00       Matches:         237
Percent Similarity:       62.14%        Conservative:    83
Best Local Similarity:    46.02%        Mismatches:      169
Query Match:              35.49%        Indels:          26
DB:                       4             Gaps:            7

US-09-806-194-1 (1-1804) x US-09-548-367D-8 (1-501)

Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84

Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129

Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla--------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189

Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAAC 249

Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309

Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369

Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429

Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489

Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549

Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609

Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669

Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720

Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780

Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840

Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900

Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
```

```
Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCCGCGCATCICTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              :::::::::      ||||||         :::|||   |||||||||,|||   |||
Db        307 LysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGlu 326

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||   ||||||           ||||||||||:::   |||||   ||||||:::
Db        327 GlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePheProValIleSerLeu 346

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||      |||  :::::::::|||||||||||||||||||||||||.|   | |
Db        347 TyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyr 366

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCC 1197
              ::::::|||:::                 :::      :::|||||||:::|||   :::|||
Db        367 LeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaValSer 386

Qy       1198 CCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTC 1257
                 |||:::         |||:::||||-|   :::|||||||||||||||||:::|:|
Db        387 GlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPhe 406

Qy       1258 GACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCT 1317
              |||||||||:::|||||||:::|||||||||   |||   ,:|
Db        407 AspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPhe 426

Qy       1318 GCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCC 1377
                    :::   :::   |||||||||   |||   |||:::
Db        427 ArgThrAlaAlaValGluGlyProPheValThrAlaAspMetGluAspCysGlyTyrAsn 446

Qy       1378 GCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGT 1437
                       |||      |||      :::::::||   :::   :::::::|||
Db        447 IleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys 466

Qy       1438 GGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT--- 1494
                 |||:::   ::::::   :::   |||::::::      :::|||||   | :
Db        467 ---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCys 485

Qy       1495 ---CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
                  |||   :::      :::      :::|:|   |||   |||
Db        486 LeuArgHisGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500
```

RESULT 11
US-09-551-853D-8
; Sequence 8, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280L
; CURRENT APPLICATION NUMBER: US/09/551,853D
; CURRENT FILING DATE: 2000-04-18
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-551-853D-8

Alignment Scores:

Pred. No.:	2.65e-95	Length:	501
Score:	1185.00	Matches:	237
Percent Similarity:	62.14%	Conservative:	83
Best Local Similarity:	46.02%	Mismatches:	169
Query Match:	35.49%	Indels:	26
DB:	4	Gaps:	7

US-09-806-194-1 (1-1804) x US-09-551-853D-8 (1-501)

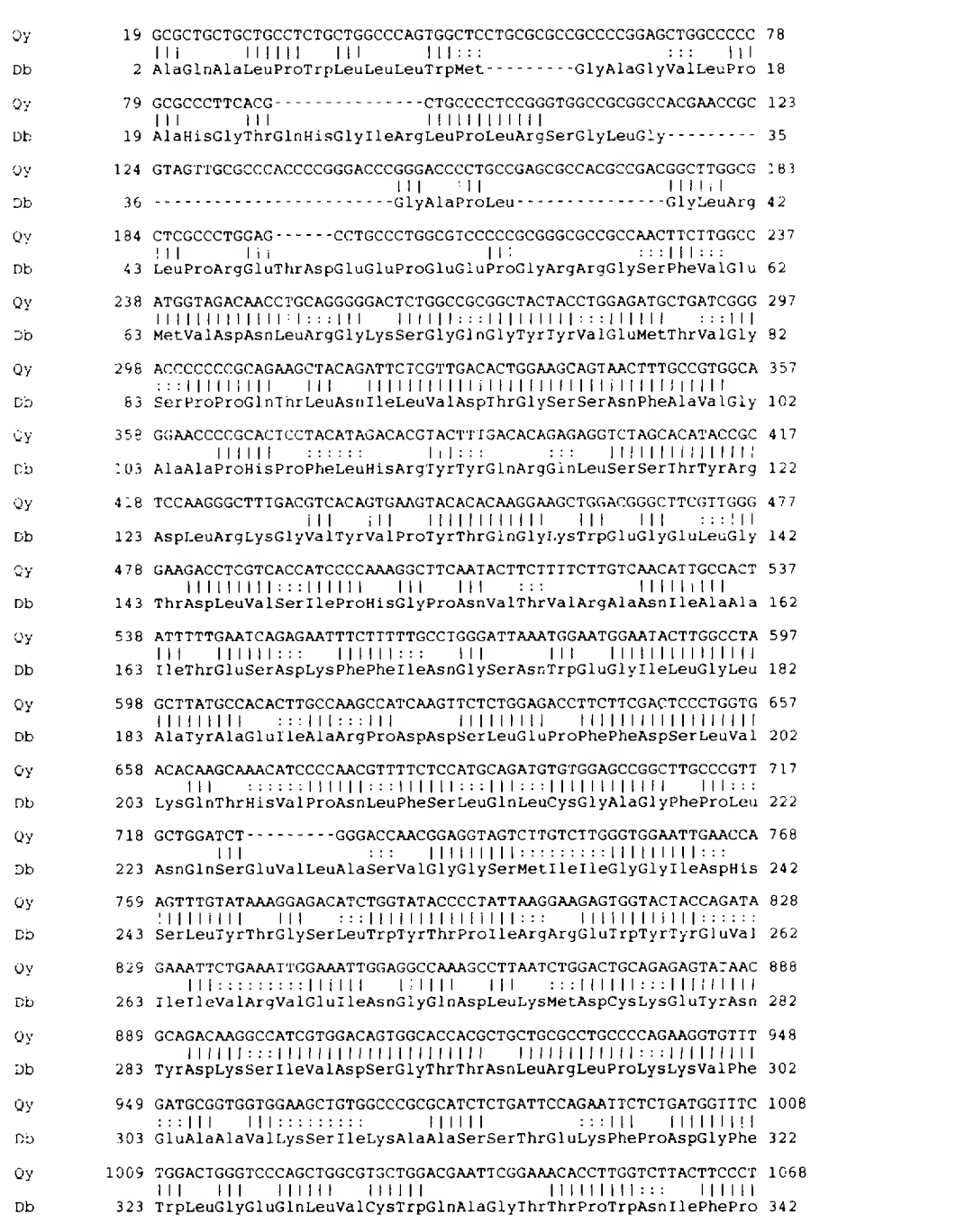
```
Qy         25 CTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCCGCGCCC 84
              :::   |||   |||         ||||||                :::   ||||||
Db          1 MetAlaProAlaLeuHisTrpLeuLeuLeuTrpValGlySerGlyMetLeuProAlaGln 20

Qy         85 TTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGCGTAGTT 129
                 |||                |||||||||||          |||
Db         21 GlyThrHisLeuGlyIleArgLeuProLeuArgSerGlyLeuAla--------------- 35

Qy        130 GCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCGCTCGCC 189
                 |||      |||         |||  |||         :::
Db         36 GlyProProLeuGlyLeuArgLeuProArgGluThrAspGluGluSer------------ 51

Qy        190 CTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTCTTGGCCATGGTAGACAAC 249
                                  |||          :::|||:::   ||||||||||||
Db         52 --------------GluGluProGlyArgArgGlySerPheValGluMetValAspAsn 66

Qy        250 CTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCGCAG 309
                 |||:::|||   |||||||:::|||||||||:::||||||   :::|||:::|||||||
Db         67 LeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGln 86

Qy        310 AAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCAC 369
                 |||   |||||||||||||||||||||||||||||||||||||          ||||||
Db         87 ThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHis 106

Qy        370 TCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTT 429
                 ::::::      |||:::      :::   ||||||||||||||
Db        107 ProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLys 126

Qy        430 GACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTC 489
                 |||  |||   ||||||||||||   |||   |||   :::|||   ||||||||
Db        127 GlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuVal 146

Qy        490 ACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCA 549
                 :::|||||||   |||   |||   :::         |||||||||   |||   ||||||
Db        147 SerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSer 166

Qy        550 GAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACA 609
                 :::   ||||||||:::   |||      |||   |||||||||||||||||||||||
Db        167 AspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGlu 186

Qy        610 CTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAAC 669
                 :::|||:::|||      ||||||||||   ||||||||||||||||||   |||   :::
Db        187 IleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHis 206

Qy        670 ATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT---------GCT 720
                 ||||||||||:::||||||:::|||:::|||:|||||||||   ||||:::
Db        207 IleProAsnIlePheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnThrGlu 226

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
                    :::   |||||||||||:::::::::|||||||||:::   |||||||||
Db        227 AlaLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThr 246

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
                 |||   :::|||||||||||||||:::   |||||||||||||:::::   |||:::::::
Db        247 GlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArg 266

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
                 :::|||||||   ||||||   |||   :::||||||:::||||||||||   |||||||:::
Db        267 ValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSer 286

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
                 ||||||||||||||||||||   ||||||||||||:::||||||||||:::|||   |||
Db        287 IleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaVal 306

Qy        961 GAAGCTGTGGCCCCGCGCATCICTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
                 :::::::::      ||||||         :::|||   |||||||||||   |||
Db        307 LysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGlu 326
```

RESULT 12
US-09-548-372D-4
; Sequence 4, Application US/09548372D
; Patent No. 6420534
; GENERAL INFORMATION:
; APPLICANT: GURNEY ET AL.
; TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: 29915/6280I
; CURRENT APPLICATION NUMBER: US/09/548,372D
; CURRENT FILING DATE: 2000-04-12
; PRIOR APPLICATION NUMBER: US 60/155,493
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 09/404,133
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: PCT/US99/20881
; PRIOR FILING DATE: 1999-09-23
; PRIOR APPLICATION NUMBER: US 60/101,594
; PRIOR FILING DATE: 1998-09-24
; NUMBER OF SEQ ID NOS: 73
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
; LENGTH: 501
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-548-372D-4

Alignment Scores:

Pred. No.:	9.81e-95	Length:	501
Score:	1178.50	Matches:	240
Percent Similarity:	62.04%	Conservative:	82
Best Local Similarity:	46.24%	Mismatches:	164
Query Match:	35.29%	Indels:	33
DB:	4	Gaps:	9

US-09-806-194-1 (1-1804) x US-09-548-372D-4 (1-501)



```
Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA 1185
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422

Qy       1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442

Qy       1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC 1425
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462

Qy       1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCGGG 1485
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnIrpArg 481

Qy       1486 TGTCAGCGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 13
US-09-548-367D-4
; Sequence 4, Application US/09548367D
; Patent No. 6440698
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280H
;  CURRENT APPLICATION NUMBER: US/09/548,367D
;  CURRENT FILING DATE:  2000-04-12
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-548-367D-4

Alignment Scores:
Pred. No.:              9.8le-95        Length:         501
Score:                  1178.50         Matches:        240
Percent Similarity:     62.04%          Conservative:   82
Best Local Similarity:  46.24%          Mismatches:     164
Query Match:            35.29%          Indels:         33
DB:                     4               Gaps:           9

US-09-806-194-1 (1-1804) x US-09-548-367D-4 (1-501)

Qy         19 GCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC 78
Db          2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet---------GlyAlaGlyValLeuPro 18

Qy         79 GCGCCCTTCACG---------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC 123
Db         19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly--------- 35

Qy        124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183
Db         36 ------------------------GlyAlaProLeu--------------GlyLeuArg 42

Qy        184 CTCGCCCTGGAG------CCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCC 237
Db         43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62

Qy        238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG 297
Db         63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly 82

Qy        298 ACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA 357
Db         83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102

Qy        358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC 417
Db        103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg 122

Qy        418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGG 477
Db        123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly 142

Qy        478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT 537
Db        143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162

Qy        538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTA 597
Db        163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182

Qy        598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG 657
Db        183 AlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal 202

Qy        658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT 717
Db        203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222

Qy        718 GCTGGATCT---------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db        223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis 242

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db        243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluVal 262

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db        283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe 322

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068
Db        323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA 1185
```

```
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422

Qy       1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442

Qy       1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTGTGGATTGTGTCCTATGCGCTC 1425
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462

Qy       1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCCTTCGGG 1485
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg 481

Qy       1486 TGTCAGCCT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1538
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 14
US-09-551-853D-4
; Sequence 4, Application US/09551853D
; Patent No. 6500667
; GENERAL INFORMATION:
;  APPLICANT: GURNEY ET AL.
;  TITLE OF INVENTION: ALZHEIMER'S DISEASE SECRETASE, APP SUBSTRATES THEREFOR AND USES
;  TITLE OF INVENTION:  THEREOF
;  FILE REFERENCE: 29915/6280L
;  CURRENT APPLICATION NUMBER: US/09/551,853D
;  CURRENT FILING DATE:  2000-04-18
;  PRIOR APPLICATION NUMBER: US 60/155,493
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 09/404,133
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: PCT/US99/20881
;  PRIOR FILING DATE: 1999-09-23
;  PRIOR APPLICATION NUMBER: US 60/101,594
;  PRIOR FILING DATE: 1998-09-24
;  NUMBER OF SEQ ID NOS: 73
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 4
;   LENGTH: 501
;   TYPE: PRT
;   ORGANISM: Homo sapiens
US-09-551-853D-4

Alignment Scores:
Pred. No.:              9.81e-95        Length:         501
Score:                  1178.50         Matches:        240
Percent Similarity:     62.04%          Conservative:   82
Best Local Similarity:  46.24%          Mismatches:     164
Query Match:            35.29%          Indels:         33
DB:                     4               Gaps:           9

US-09-806-194-1 (1-1804) x US-09-551-853D-4 (1-501)

Qy         19 GCGCTGCTGCTGCCTCTGCTCGCCCAGTGGCTCCTGCGCGCCGCCCCGGAGCTGGCCCCC 78
Db          2 AlaGlnAlaLeuProTrpLeuLeuLeuTrpMet---------GlyAlaGlyValLeuPro 18

Qy         79 GCGCCCTTCACG--------------CTGCCCCTCCGGGTGGCCGCGGCCACGAACCGC 123
Db         19 AlaHisGlyThrGlnHisGlyIleArgLeuProLeuArgSerGlyLeuGly--------- 35

Qy        124 GTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTGGCG 183

Db         36 ----------------------GlyAlaProLeu---------------GlyLeuArg 42

Qy        184 CTCGCCCTGGAG------CCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCC 237
Db         43 LeuProArgGluThrAspGluGluProGluGluProGlyArgArgGlySerPheValGlu 62

Qy        238 ATGGTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGG 297
Db         63 MetValAspAsnLeuArgGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGly 82

Qy        298 ACCCCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCA 357
Db         83 SerProProGlnThrLeuAsnIleLeuValAspThrGlySerSerAsnPheAlaValGly 102

Qy        358 GGAACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGC 417
Db        103 AlaAlaProHisProPheLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArg 122

Qy        418 TCCAAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGG 477
Db        123 AspLeuArgLysGlyValTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGly 142

Qy        478 GAAGACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACT 537
Db        143 ThrAspLeuValSerIleProHisGlyProAsnValThrValArgAlaAsnIleAlaAla 162

Qy        538 ATTTTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTA 597
Db        163 IleThrGluSerAspLysPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeu 182

Qy        598 GCTTATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTG 657
Db        183 AlaTyrAlaGluIleAlaArgProAspAspSerLeuGluProPhePheAspSerLeuVal 202

Qy        658 ACACAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTT 717
Db        203 LysGlnThrHisValProAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeu 222

Qy        718 GCTGGATCT---------GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCA 768
Db        223 AsnGlnSerGluValLeuAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHis 242

Qy        769 AGTTTGTATAAAGGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATA 828
Db        243 SerLeuTyrThrGlySerLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluVal 262

Qy        829 GAAATTCTGAAATTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAAC 888
Db        263 IleIleValArgValGluIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsn 282

Qy        889 GCAGACAAGGCCATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTT 948
Db        283 TyrAspLysSerIleValAspSerGlyThrThrAsnLeuArgLeuProLysLysValPhe 302

Qy        949 GATGCGGTGGTGGAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTC 1008
Db        303 GluAlaAlaValLysSerIleLysAlaAlaSerSerThrGluLysPheProAspGlyPhe 322

Qy       1009 TGGACTGGGTCCCAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCT 1068
Db        323 TrpLeuGlyGluGlnLeuValCysTrpGlnAlaGlyThrThrProTrpAsnIlePhePro 342

Qy       1069 AAAATCTCCATCTACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTG 1128
Db        343 ValIleSerLeuTyrLeuMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeu 362

Qy       1129 CCTCAGCTTTACATTCAGCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGA 1185
Db        363 ProGlnGlnTyrLeuArgProValGluAspValAlaThrSerGlnAspAspCysTyrLys 382

Qy       1186 TTCGGCATTTCCCCATCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTC 1245
```

```
Db        383 PheAlaIleSerGlnSerSerThrGlyThrValMetGlyAlaValIleMetGluGlyPhe 402

Qy       1246 TACGTCATCTTCGACAGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAA 1305
              ||||||:::|||||||||||||:::|||||:::||||.|||!   |||   |||
Db        403 TyrValValPheAspArgAlaArgLysArgIleGlyPheAlaValSerAlaCysHisVal 422

Qy       1306 ATTGCAGGTGCTGCAGTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGC 1365
                        :::   :::   ||||||||||   |||   :|||:::
Db        423 HisAspGluPheArgThrAlaAlaValGluGlyProPheValThrLeuAspMetGluAsp 442

Qy       1366 AACTGTGTCCCCGCTCAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTC 1425
                                  |||      |||      :::::::|,   :::
Db        443 CysGlyTyrAsnIleProGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMet 462

Qy       1426 ATGAGCGTCTGTGGAGCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGC 1485
                :::::::|||   |||:::   ::::::   :::   ||,:::::::      :::|.|
Db        463 AlaAlaIleCys---AlaLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArg 481

Qy       1486 TGTCAGCGT------CGCCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
              |||   |||      |||   :::      :::      :::|||   |||   |||
Db        482 CysLeuArgCysLeuArgGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 500

RESULT 15
US-09-604-608-3
; Sequence 3, Application US/09604608
; Patent No. 6545127
; GENERAL INFORMATION:
;  APPLICANT:  Tang, Jordan J.N.
;  APPLICANT:  Lin, Xinli
;  APPLICANT:  Koelsch, Gerald
;  TITLE OF INVENTION:  Catalytically Active Recombinant Memapsin and Methods
;  TITLE OF INVENTION:  of Use Thereof
;  FILE REFERENCE:  OMRF 179
;  CURRENT APPLICATION NUMBER: US/09/604,608
;  CURRENT FILING DATE:  2000-06-27
;  PRIOR APPLICATION NUMBER:  60/141,363
;  PRIOR FILING DATE:  1999-06-28
;  PRIOR APPLICATION NUMBER:  60/168,060
;  PRIOR FILING DATE:  1999-11-30
;  PRIOR APPLICATION NUMBER:  60/177,836
;  PRIOR FILING DATE:  2000-01-25
;  PRIOR APPLICATION NUMBER:  60/178,368
;  PRIOR FILING DATE:  2000-01-27
;  PRIOR APPLICATION NUMBER:  60/210,292
;  PRIOR FILING DATE:  2000-06-08
;  NUMBER OF SEQ ID NOS:  31
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 3
;  LENGTH: 503
;  TYPE: PRT
;  ORGANISM: Homo sapiens
;  FEATURE:
;  OTHER INFORMATION: Pro-memapsin 2
;  OTHER INFORMATION: Amino Acids 1-15 are vector derived residues
;  OTHER INFORMATION: Amino Acids 16-64 are a putative pro peptide
;  OTHER INFORMATION: Amino Acids 1-13 are the T7 promoter
;  OTHER INFORMATION: Amino Acids 16-456 are Pro-memapsin 2-T1
;  OTHER INFORMATION: Amino Acids 16-421 are Promemapsin 2-T2
US-09-604-608-3

Alignment Scores:
Pred. No.:              1.8e-94         Length:         503
Score:                  1175.50         Matches:        235
Percent Similarity:     63.97%          Conservative:   81
Best Local Similarity:  47.57%          Mismatches:     160
Query Match:            35.21%          Indels:         18
DB:                     4               Gaps:           6

US-09-806-194-1 (1-1804) x US-09-604-608-3 (1-503)

Qy        105 GGCCGCGGCCACGAACCGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCG 164
              |||||||||                  ||||||:::       |||
Db         11 GlyArgGlySerMetAlaGlyValLeuProAlaHisGlyThrGlnHisGlyIleArgLeu 30

Qy        165 CCA---CGCCGACGGCTTGGCGCTCGCCCTGGAGCCTGC---------- 200
              |||   |||      ||||||      |||
Db         31 ProLeuArgSerGlyLeuGlyGlyAlaProLeuGlyLeuArgLeuProArgGluThrAsp 50

Qy        201 ------CCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATGGTAGACAACCTGCA 254
                   |||      |||      |||   :::|||:::   |||||||||||||||::
Db         51 GluGluProGluGluProGlyArgArg-GlySerPheValGluMetValAspAsnLeuAr 70

Qy        255 GGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACCCCCCCCGCAGAAGCT 314
              :|||   |||||||:::|||||||||:::||||||   :::|||:::|||||||||   ||
Db         70 gGlyLysSerGlyGlnGlyTyrTyrValGluMetThrValGlySerProProGlnThrLe 90

Qy        315 ACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGAACCCCGCACTCCTA 374
              |   |||||||||||||||||||||||||||||||||||||||||         ||||||   ::
Db         90 uAsnIleLeuValAspThrGlySerSerAsnPheAlaValGlyAlaAlaProHisProPh 110

Qy        375 CATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCCAAGGGCTTTGACGT 434
              ::::      |||:::      :::   ||||||||||||||||               ||
Db        110 eLeuHisArgTyrTyrGlnArgGlnLeuSerSerThrTyrArgAspLeuArgLysGlyVa 130

Qy        435 CACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAAGACCTCGTCACCAT 494
              |   |||   ||||||||||||   |||   |||   :::|||   |||||||||:::||
Db        130 lTyrValProTyrThrGlnGlyLysTrpGluGlyGluLeuGlyThrAspLeuValSerIl 150

Qy        495 CCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATTTTTGAATCAGAGAA 554
              ||||   |||   |||   :::          !||||||||   |||   |||||||:::
Db        150 eProHisGlyProAsnValThrValArgAlaAsnIleAlaAlaIleThrGluSerAspLy 170

Qy        555 TTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCTTATGCCACACTTGC 614
               |||||:::   |||      |||   ||||||||||||||||||||||||   :::||
Db        170 sPhePheIleAsnGlySerAsnTrpGluGlyIleLeuGlyLeuAlaTyrAlaGluIleAl 190

Qy        615 CAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACACAAGCAAACATCCC 674
              |:::||      |||||||||   ||||||||||||||||||   |||   :::::::||
Db        190 aArgProAspAspSerLeuGluProPhePheAspSerLeuValLysGlnThrHisValPr 210

Qy        675 CAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCCGTTGCTGGATCT-------- 726
              ||||:::|||||||:::|||:::||||||||||||   |||:::      |||
Db        210 oAsnLeuPheSerLeuGlnLeuCysGlyAlaGlyPheProLeuAsnGlnSerGluValLe 230

Qy        727 -GGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAAGGAGA 785
                  :::   |||||||||:::::::::||:|||||||:::   ||||||||   |||
Db        230 uAlaSerValGlyGlySerMetIleIleGlyGlyIleAspHisSerLeuTyrThrGlySe 250

Qy        786 CATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAATTGGA 845
              :::||||:|||||||||:::   |||||||||||||:::::   ||:::::::::||
Db        250 rLeuTrpTyrThrProIleArgArgGluTrpTyrTyrGluValIleIleValArgValGl 270

Qy        846 AATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCCATCGT 905
              ||||   ||||||   |||   :::|||||:::|||||||||   ||||||:::||||
Db        270 uIleAsnGlyGlnAspLeuLysMetAspCysLysGluTyrAsnTyrAspLysSerIleVa 290

Qy        906 GGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTGGAAGC 965
              ||||:|||||||||||   |||||||||||:::|||||||||:::|||   |||:::::
Db        290 lAspSerGlyThrThrAsnLeuArgLeuProLysLysValPheGluAlaAlaValLysSe 310

Qy        966 TGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCCCAGCT 1025
              ::::      ||||||          :::|||   ||||||||||||   |||   |||||
Db        310 rIleLysAlaAlaSerSerThrGluLysPheProAspGlyPheTrpLeuGlyGluGlnLe 330

Qy       1026 GGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATCTACCT 1085
              |   ||||||            |||||||||:::   ||||||   ||||||:::|||||
Db        330 uValCysTrpGlnAlaGlyThrThrProTrpAsnIleProValIleSerLeuTyrLe 350

Qy       1086 GAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTACATTCA 1145
              |      |||   :::::::::|||||||||||||||||||||||||||||   |||:::::
Db        350 uMetGlyGluValThrAsnGlnSerPheArgIleThrIleLeuProGlnGlnTyrLeuAr 370

Qy       1146 GCCCATGATGGGGGCCGGCCTGAATTAT---GAATGTTACCGATTCGGCATTTCCCCATC 1202
              :|||:::                :::      :::||||||:::|||   ||||||   ||
Db        370 gProValGluAspValAlaThrSerGlnAspAspCysTyrLysPheAlaIleSerGlnSe 390
```

```
Qy       1203 CACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGACAG 1262
Db        390 rSerThrGlyThrValMetGlyAlaValIleMetGluGlyPheTyrValValPheAspAr 410

Qy       1263 AGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCAGT 1322
Db        410 gAlaArgLysArgIleGlyPheAlaValSerAlaCysHisValHisAspGluPheArgTh 430

Qy       1323 GTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCTCA 1382
Db        430 rAlaAlaValGluGlyProPheValThrLeuAspMetGluAspCysGlyTyrAsnIlePr 450

Qy       1383 GTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGAGC 1442
Db        450 oGlnThrAspGluSerThrLeuMetThrIleAlaTyrValMetAlaAlaIleCys---Al 469

Qy       1443 CATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGT------CG 1496
Db        469 aLeuPheMetLeuProLeuCysLeuMetValCysGlnTrpArgCysLeuArgCysLeuAr 489

Qy       1497 CCCCCGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTG 1536
Db        489 gGlnGlnHisAspAspPheAlaAspAspIleSerLeuLeu 502
```

Search completed: September 25, 2003, 11:10:56
Job time : 62.5 secs

GenCore version 5.1.6

OM nucleic - protein search, using frame_plus_n2p model

Run on: September 25, 2003, 10:19:24 ; Search time 114.5 Seconds
(without alignments)
5001.620 Million cell updates/sec

Title: US-09-806-194-1
Perfect score: 3339
Sequence: 1 atgggcgcactggccccgggc.....aaaccaaaaaaaaaaaaaaa 1804

Scoring table: BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 1107863 seqs, 158726573 residues

Total number of hits satisfying chosen parameters: 2215726

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODEL=frame+_n2p.model -DEV=xlp
-Q=/cgn2_1/USPTO_spool/US09806194/runat_25092003_095453_10998/app_query.fasta_1.1991
-DB=A_Geneseq_19Jun03 -QFMT=fastan -SUFFIX=rag -MINMATCH=0.1 -LOOPCL=0
-LOOPEXT=0 -UNITS=bits -START=1 -END=-1 -MATRIX=blosum62 -TRANS=human40.cdi
-LIST=45 -DOCALIGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTFMT=pto -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-USER=US09806194_@CGN_1_1_158_@runat_25092003_095453_10998 -NCPU=6 -ICPU=3
-NO_MMAP -LARGEQUERY -NEG_SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV_TIMEOUT=120 -WARN_TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : A_Geneseq_19Jun03:*
1: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1980.DAT:*
2: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1981.DAT:*
3: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1982.DAT:*
4: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1983.DAT:*
5: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1984.DAT:*
6: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1985.DAT:*
7: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1986.DAT:*
8: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1987.DAT:*
9: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1988.DAT:*
10: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1989.DAT:*
11: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1990.DAT:*
12: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1991.DAT:*
13: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1992.DAT:*
14: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1993.DAT:*
15: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1994.DAT:*
16: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1995.DAT:*
17: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1996.DAT:*
18: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1997.DAT:*
19: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1998.DAT:*
20: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA1999.DAT:*
21: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2000.DAT:*
22: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2001.DAT:*
23: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2002.DAT:*
24: /SIDS1/gcgdata/geneseq/geneseqp-embl/AA2003.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	2687	80.5	518	19	AAW61362	Aspartic proteinas
2	2687	80.5	518	20	AAY41714	Human PRO852 prote
3	2687	80.5	518	20	AAY22239	Human CSP56, aspar
4	2687	80.5	518	20	AAY13799	Human aspartyl pro
5	2687	80.5	518	21	AAB44270	Human PRO852 (UNQ4
6	2687	80.5	518	21	AAY88424	Human aspartyl pro
7	2687	80.5	518	22	AAE10628	Human aspartyl pro
8	2687	80.5	518	22	AAE10656	Human-Asp 1 protei
9	2687	80.5	518	22	AAU29059	Human PRO polypept
10	2687	80.5	518	22	AAE06858	Human aspartyl pro
11	2687	80.5	518	22	AAU06602	Human Aspartyl pro
12	2687	80.5	518	22	AAU07201	Human aspartyl pro
13	2687	80.5	518	22	AAE02580	Human aspartyl pro
14	2687	80.5	518	22	AAE02608	Human Aspartyl pro
15	2687	80.5	518	23	ABB78589	Human Asp-1 protei
16	2687	80.5	518	23	ABB78617	Human Asp-1deltaTM
17	2687	80.5	518	23	ABB06531	Human aspartyl pro
18	2687	80.5	518	23	ABB07453	Human BACE2 amino
19	2687	80.5	518	24	ABU71147	Human PRO852 prote
20	2687	80.5	518	24	ABU65604	Human secreted/tra
21	2687	80.5	518	24	ABU65937	Novel human secret
22	2687	80.5	518	24	ABU67441	Human secreted/tra
23	2687	80.5	518	24	ABU61100	Human PRO852 polyp
24	2687	80.5	518	24	ABU65299	Human PRO polypept
25	2687	80.5	518	24	ABU58435	Human PRO polypept
26	2687	80.5	518	24	ABU55971	Human secreted/tra
27	2687	80.5	518	24	ABU56966	Human PRO polypept
28	2687	80.5	518	24	ABU10545	Human secreted/tra
29	2506	75.1	481	22	AAG75592	Human colon cancer
30	2437	73.0	475	22	AAE10657	Secreted recombina
31	2437	73.0	475	22	AAE02609	Human secreted asp
32	2437	73.0	475	23	ABB78618	Secreted recombina
33	2395	71.7	514	22	AAB84204	Amino acid sequenc
34	2293	68.7	439	23	ABB90365	Human polypeptide
35	2207	66.1	423	22	AAE04796	Human aspartyl pro
36	2184	65.4	423	22	AAB88479	Human membrane or
37	2124	63.6	413	22	AAE10658	Acid-processed hu-
38	2124	63.6	413	22	AAE02610	Human acid-process
39	2124	63.6	413	23	ABB78619	Asp-1deltaTM(his)6
40	1794	53.7	355	22	AAB93925	Human protein sequ
41	1187	35.5	501	21	AAY94769	Rat beta-secretase
42	1186.5	35.5	501	22	AAB84948	Mouse aspartic sec
43	1185	35.5	501	21	AAY94768	Murine beta-secret
44	1185	35.5	501	21	AAY88427	Murine aspartyl pr
45	1185	35.5	501	22	AAE10631	Murine aspartyl pr

ALIGNMENTS

RESULT 1
AAW61362
ID AAW61362 standard; Protein; 518 AA.
XX
AC AAW61362;
XX
DT 25-MAR-2003 (updated)
DT 25-SEP-1998 (first entry)
XX
DE Aspartic proteinase ASP1.
XX
KW ASP1; Aspartic proteinase; Alzheimer's disease; cancer; melanoma.
XX
OS Homo sapiens.
XX
PN EP848062-A2.
XX
PD 17-JUN-1998.

```
XX
PF   01-DEC-1997;   97EP-0309648.
XX
PR   14-DEC-1996;   96GB-0026022.
PR   06-OCT-1997;   97US-0999723.
XX
PA   (SMIK ) SMITHKLINE BEECHAM CORP.
PA   (SMIK ) SMITHKLINE BEECHAM PLC.
XX
PI   Chapman CG,  Evans JR,  Powell DJ,  Southan C;
XX
DR   WPI; 1998-314477/28.
DR   N-PSDB; AAV27962.
XX
PT   New isolated polynucleotide encodes Aspartic protease polypeptide -
PT   used to diagnosis, treat and vaccinate against Alzheimer's disease,
PT   cancer and melanoma
XX
PS   Claim 11; Page 7; 19pp; English.
XX
CC   The human ASP1 protein is structurally related to other proteins of the
CC   Aspartic proteinase family.  ASP 1 polypeptides and polynucleotides can
CC   be used to diagnosis, treat and vaccinate against Alzheimer's disease,
CC   cancer and melanoma.
CC   (Updated on 25-MAR-2003 to correct PR field.)
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     19              Gaps:           0

US-09-806-194-1 (1-1804) x AAW61362 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 2
AAY41714
ID   AAY41714 standard; Protein; 518 AA.
XX
```

```
AC   AAY41714;
XX
DT   07-DEC-1999  (first entry)
XX
DE   Human PRO852 protein sequence.
XX
KW   Human; PRO; EST; expressed sequence tag; PCR primer; hybridisation;
KW   probe; blood coagulation disorder; cancer; cellular adhesion disorder;
KW   secreted protein; transmembrane protein.
XX
OS   Homo sapiens.
XX
PN   WO9946281-A2.
XX
PD   16-SEP-1999.
XX
PF   08-MAR-1999;   99WO-US05028.
XX
PR   10-MAR-1998;   98US-0077450.
PR   11-MAR-1998;   98US-0077632.
PR   11-MAR-1998;   98US-0077641.
PR   11-MAR-1998;   98US-0077649.
PR   12-MAR-1998;   98US-0077791.
PR   13-MAR-1998;   98US-0078004.
PR   17-MAR-1998;   98US-0040220.
PR   20-MAR-1998;   98US-0078886.
PR   20-MAR-1998;   98US-0078910.
PR   20-MAR-1998;   98US-0078936.
PR   20-MAR-1998;   98US-0078939.
PR   25-MAR-1998;   98US-0079294.
PR   26-MAR-1998;   98US-0079656.
PR   27-MAR-1998;   98US-0079663.
PR   27-MAR-1998;   98US-0079664.
PR   27-MAR-1998;   98US-0079689.
PR   27-MAR-1998;   98US-0079728.
PR   27-MAR-1998;   98US-0079786.
PR   30-MAR-1998;   98US-0079920.
PR   30-MAR-1998;   98US-0079923.
PR   31-MAR-1998;   98US-0080105.
PR   31-MAR-1998;   98US-0080107.
PR   31-MAR-1998;   98US-0080165.
PR   31-MAR-1998;   98US-0080194.
PR   01-APR-1998;   98US-0080327.
PR   01-APR-1998;   98US-0080328.
PR   01-APR-1998;   98US-0080333.
PR   01-APR-1998;   98US-0080334.
PR   08-APR-1998;   98US-0081049.
PR   08-APR-1998;   98US-0081070.
PR   08-APR-1998;   98US-0081071.
PR   09-APR-1998;   98US-0081195.
PR   09-APR-1998;   98US-0081203.
PR   09-APR-1998;   98US-0081229.
PR   15-APR-1998;   98US-0081817.
PR   15-APR-1998;   98US-0081838.
PR   15-APR-1998;   98US-0081952.
PR   15-APR-1998;   98US-0081955.
PR   21-APR-1998;   98US-0082568.
PR   21-APR-1998;   98US-0082569.
PR   22-APR-1998;   98US-0082700.
PR   22-APR-1998;   98US-0082704.
PR   22-APR-1998;   98US-0082804.
PR   23-APR-1998;   98US-0082767.
PR   23-APR-1998;   98US-0082796.
PR   27-APR-1998;   98US-0083336.
PR   28-APR-1998;   98US-0083322.
PR   29-APR-1998;   98US-0083392.
PR   29-APR-1998;   98US-0083495.
PR   29-APR-1998;   98US-0083496.
PR   29-APR-1998;   98US-0083499.
PR   29-APR-1998;   98US-0083500.
PR   29-APR-1998;   98US-0083545.
PR   29-APR-1998;   98US-0083554.
PR   29-APR-1998;   98US-0083558.
PR   29-APR-1998;   98US-0083559.
PR   30-APR-1998;   98US-0083742.
PR   05-MAY-1998;   98US-0084366.
PR   06-MAY-1998;   98US-0084414.
PR   06-MAY-1998;   98US-0084441.
PR   07-MAY-1998;   98US-0084598.
PR   07-MAY-1998;   98US-0084600.
PR   07-MAY-1998;   98US-0084627.
PR   07-MAY-1998;   98US-0084637.
PR   07-MAY-1998;   98US-0084639.
PR   07-MAY-1998;   98US-0084640.
PR   07-MAY-1998;   98US-0084643.
PR   13-MAY-1998;   98US-0085323.
PR   13-MAY-1998;   98US-0085338.
PR   13-MAY-1998;   98US-0085339.
PR   15-MAY-1998;   98US-0085573.
PR   15-MAY-1998;   98US-0085579.
PR   15-MAY-1998;   98US-0085580.
PR   15-MAY-1998;   98US-0085582.
PR   15-MAY-1998;   98US-0085689.
PR   15-MAY-1998;   98US-0085697.
PR   15-MAY-1998;   98US-0085700.
PR   15-MAY-1998;   98US-0085704.
PR   18-MAY-1998;   98US-0086023.
PR   22-MAY-1998;   98US-0086392.
PR   22-MAY-1998;   98US-0086414.
PR   22-MAY-1998;   98US-0086430.
PR   22-MAY-1998;   98US-0086486.
PR   28-MAY-1998;   98US-0087098.
PR   28-MAY-1998;   98US-0087106.
PR   28-MAY-1998;   98US-0087208.
PR   30-JUL-1998;   98US-0094651.
PR   11-SEP-1998;   98US-0100038.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Wood WI, Goddard A, Gurney A, Yuan J, Baker KP, Chen J;
XX
DR   WPI; 1999-551358/46.
DR   N-PSDB; AAZ34056.
XX
PT   New secreted and transmembrane polypeptides and their polynucleotides,
PT   useful for treating blood coagulation disorders, cancers and cellular
PT   adhesion disorders -
XX
PS   Claim 12; Fig 73; 530pp; English.
XX
CC   The present invention describes secreted and transmembrane polypeptides
CC   and their polynucleotides. The nucleotide sequences are useful as
CC   sources of probes, primers, for chromosome mapping, and for generation
CC   of antisense sequences. They can also be used to create transgenic
CC   animals. The proteins can be used to treat a variety of diseases and
CC   disorders, depending on their function. Diseases that may be treated
CC   include blood coagulation disorders, cancers and cellular adhesion
CC   disorders. They may also be used to raise antibodies. AAZ33891 to
CC   AAZ34338, and AAY41685 to AAY41774 represent polynucleotide and
CC   polypeptide sequence given in the exemplification of the present
CC   invention.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:                4.01e-213       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       20              Gaps:            0

US-09-806-194-1 (1-1804) x AAY41714 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||.
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCG:TGCT 720
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAAA 1554
               ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

```
RESULT 3
AAY22239
ID   AAY22239 standard; Protein; 518 AA.
XX
AC   AAY22239;
XX
DT   20-SEP-1999  (first entry)
XX
DE   Human CSP56, aspartyl-type protease, protein sequence.
XX
KW   Metastatic marker protein; human; cancer metastasis; breast cancer;
KW   colon cancer; diagnosis; therapy; tumour; metastatic potential;
KW   CSP56; aspartyl-type protease.
XX
OS   Homo sapiens.
XX
PN   WO9934004-A2.
XX
PD   08-JUL-1999.
XX
PF   24-DEC-1998;   98WO-US27608.
XX
PR   31-DEC-1997;   97US-0070112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Giese K,  Xin H;
XX
DR   WPI; 1999-430248/36.
DR   N-PSDB; AAX84708.
XX
PT   New polynucleotides associated with cancer metastasis
XX
PS   Claim 4; Page 78-80; 80pp; English.
XX
CC   This sequence represents a polypeptide of the invention, and is
CC   an aspartyl-type protease, designated CSP56. The polynucleotides (PNs) of
CC   the invention encode metastatic marker protein variants. The PNs and
CC   polypeptides can be used as markers for cancer metastasis. The products
CC   can be used for identifying metastatic tissue or metastatic potential of
CC   a tissue, e.g. breast or colon tissue. They can also be used for
CC   screening test compounds for the ability to suppress the metastatic
CC   potential of a tumour. The products can be used for developing products
CC   for the therapy of cancers, particularly breast or colon cancer.
```

```
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     20              Gaps:           0

US-09-806-194-1 (1-1804) x AAY22239 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTCAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCAGTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 4
AAY13799
ID   AAY13799 standard; Protein; 518 AA.
XX
AC   AAY13799;
XX
DT   21-SEP-1999  (first entry)
XX
DE   Human aspartyl protease, CSP56.
XX
KW   CSP56; human; aspartyl protease; diagnosis; neoplasia; tumour;
KW   breast tumour; colon tumour.
XX
OS   Homo sapiens.
XX
PN   WO9933963-A1.
XX
PD   08-JUL-1999.
XX
PF   14-DEC-1998;   98WO-US26547.
XX
PR   31-DEC-1997;   97US-0070112.
XX
PA   (CHIR ) CHIRON CORP.
XX
PI   Giese KW,  Xin H;
XX
DR   WPI; 1999-430240/36.
DR   N-PSDB; AAX89297.
```

```
XX
PT   Human CSP56 protein for diagnosis of neoplasia
XX
PS   Claim 2; Fig 2A; 51pp; English.
XX
CC   This represents a human CSP56 protein, a novel aspartyl protease. The
CC   CSP56 protein can be used in methods for diagnosing neoplasia, for
CC   determining the metastatic potential of a tumour, and for screening test
CC   compounds for the ability to suppress the metastatic potential of a
CC   tumour. The tumours are preferably from breast or colon.
XX
SQ   Sequence   518 AA;
```

Alignment Scores:

Pred. No.:	4.01e-213	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	20	Gaps:	0

US-09-806-194-1 (1-1804) x AAY13799 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||

Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

```
RESULT 5
AAB44270
ID   AAB44270 standard; Protein; 518 AA.
XX
AC   AAB44270;
XX
DT   08-FEB-2001  (first entry)
XX
DE   Human PRO852 (UNQ418) protein sequence SEQ ID NO:196.
XX
KW   Human; secreted protein; transmembrane protein; PRO; EST; cytostatic;
KW   expressed sequence tag; detection; cancer.
XX
OS   Homo sapiens.
XX
PN   WO200053756-A2.
XX
PD   14-SEP-2000.
XX
```

```
PF   18-FEB-2000; 2000WO-US04341.
XX
PR   08-MAR-1999;   99WO-US05028.
PR   12-MAR-1999;   99US-0123957.
PR   29-MAR-1999;   99US-0126773.
PR   21-APR-1999;   99US-0130232.
PR   28-APR-1999;   99US-0131445.
PR   14-MAY-1999;   99US-0134287.
PR   23-JUN-1999;   99US-0141037.
PR   26-JUL-1999;   99US-0145698.
PR   29-OCT-1999;   99US-0162506.
PR   30-NOV-1999;   99WO-US28313.
PR   02-DEC-1999;   99WO-US28551.
PR   02-DEC-1999;   99WO-US28565.
PR   16-DEC-1999;   99WO-US30095.
PR   30-DEC-1999;   99WO-US31243.
PR   30-DEC-1999;   99WO-US31274.
PR   05-JAN-2000; 2000WO-US00219.
PR   06-JAN-2000; 2000WO-US00277.
PR   06-JAN-2000; 2000WO-US00376.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Ashkenazi AJ,  Baker KP,  Botstein D,  Desnoyers L,  Eaton DL;
PI   Ferrara N,  Filvaroff E,  Fong S,  Gao W,  Gerber H,  Gerritsen ME;
PI   Goddard A,  Godowski PJ,  Grimaldi CJ,  Gurney AL,  Hillan KJ;
PI   Kljavin IJ,  Kuo SS,  Napier MA,  Pan J,  Paoni NF,  Roy MA;
PI   Shelton DL,  Stewart TA,  Tumas D,  Williams PM,  Wood WI;
XX
DR   WPI; 2000-611443/58.
DR   N-PSDB; AAC78500.
XX
PT   Novel PRO polypeptides and polynucleotides used in detection methods,
PT   to target bioactive molecules to specific cells, and to modulate
PT   cellular activities -
XX
PS   Claim 12; Fig 73; 636pp; English.
XX
CC   AAC78458 to AAC78599 represent polynucleotide and EST (expressed
CC   sequence tag) sequences which encode secreted or transmembrane PRO
CC   polypeptides. The PRO polynucleotides and polypeptides have cytostatic
CC   activity. The polynucleotides and polypeptides can be used for detecting
CC   the presence of PRO polypeptides in samples, for linking bioactive
CC   molecules to cells and for modulating biological activities of cells,
CC   using the polypeptides for specific targeting. The polypeptide targeting
CC   can be used to kill the target cells, e.g. for the treatment of cancers.
CC   The polypeptide pairs provide specific targeting of bioactive molecules
CC   to cells. AAC78600 to AAC78987 represent PCR primers and probes used in
CC   the isolation of the PRO polynucleotide sequences.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213      Length:          518
Score:                  2687.00        Matches:         518
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            80.47%         Indels:          0
DB:                     21             Gaps:            0

US-09-806-194-1 (1-1804) x AAB44270 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTGCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              |||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTICTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGCCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGITTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420
```

```
Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
             ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 6
AAY88424
ID   AAY88424 standard; Protein; 518 AA.
XX
AC   AAY88424;
XX
DT   03-AUG-2000  (first entry)
XX
DE   Human aspartyl protease 1 (Aspl) amino acid sequence.
XX
KW   Aspartyl protease; aspartase; amyloid precursor protein; APP; Asp 1;
KW   Alzheimer's disease; beta secretase site.
XX
OS   Homo sapiens.
XX
PN   WO200017369-A2.
XX
PD   30-MAR-2000.
XX
PF   23-SEP-1999;   99WO-US20881.
XX
PR   24-SEP-1998;   98US-0101594.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney ME, Bienkowski MJ, Heinrikson RL, Parodi LA, Yan R;
XX
DR   WPI; 2000-303209/26.
DR   N-PSDB; AAA15661.
XX
PT   New enzyme designated human aspartase usetul in research into
PT   Alzheimer's Disease is capable of cleaving amyloid protein precursor at
PT   the beta secretase site to produce amyloid beta peptide -
XX
PS   Claim 54; Fig 1; 183pp; English.
XX
CC   This sequence represents the human aspartyl protease amino acid sequence.
CC   The invention relates to a protease capable of cleaving the beta
CC   secretase site of amyloid precursor protein (APP). The protease contains
CC   a sequence encoding the amino acid sequence DTG and a sequence encoding
CC   DSG or DTG separated by 100-300 amino acids. When mutated the APP gene
CC   causes an autosomal dominant form of Alzheimer's disease. APP localises
CC   to the cell surface membrane and have a single C-terminal transmembrane
CC   domain. Proteolytic processing of APP produces the amyloid beta protein,
CC   which is possibly very important in Alzheimer's disease. The invention
CC   includes a nucleotide sequence encoding the protease, a vector containing
CC   the nucleotide sequence, and a cell line comprising the vector. Methods
CC   for screening for inhibitors of beta secretase activity are also given in
CC   the invention. The human aspartase protein and nucleotide sequences and
CC   the methods for identifying inhibitors of the protease, are useful in the
CC   treatment of and research in to Alzheimer's disease.
XX
SQ   Sequence   518 AA;
```

```
Alignment Scores:
Pred. No.:                4.01e-213      Length:          518
Score:                    2687.00        Matches:         518
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              80.47%         Indels:          0
DB:                       21             Gaps:            0

US-09-806-194-1 (1-1804) x AAY88424 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
```

```
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 7
AAE10628
ID   AAE10628 standard; Protein; 518 AA.
XX
AC   AAE10628;
XX
DT   10-DEC-2001  (first entry)
XX
DE   Human aspartyl protease 1 (hu-Aspl) protein.
XX
KW   Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
KW   Alzheimer's disease; AD; dementia;  neurofibrillary tangle; gliosis;
KW   amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective;
KW   chromosome 21.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..20
FT                   /label= Signal_peptide
FT   Protein         21..518
FT                   /note= "Mature human aspartyl protease 1"
FT   Domain          469..492
FT                   /label= Transmembrane_domain
XX
PN   GB2357767-A.
XX
PD   04-JUL-2001.
XX
PF   22-SEP-2000; 2000GB-0023315.
XX
PR   23-SEP-1999;   99US-0155493.
PR   23-SEP-1999;   99US-0404133.
PR   23-SEP-1999;   99WO-US20881.
PR   13-OCT-1999;   99US-0416901.
PR   06-DEC-1999;   99US-0169232.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Bienkowksi MJ,  Gurney M;
XX
DR   WPI; 2001-444208/48.
DR   N-PSDB; AAD17864.
XX
PT   Polypeptide comprising fragments of human aspartyl protease with
PT   amyloid precursor protein processing activity and alpha-secretase
PT   activity, for identifying modulators useful in treating Alzheimer's
PT   disease -
XX
PS   Claim 36; Fig 1; 187pp; English.
XX
CC   The patent discloses human aspartyl protease 1 (hu-Aspl) or modified
CC   Aspl proteins which lack transmembrane domain or amino terminal
CC   domain or cytoplasmic domain and retains alpha-secretase activity
CC   and amyloid protein precursor (APP) processing activity. The proteins
CC   of the invention are useful for assaying hu-Aspl alpha-secretase
CC   activity, which in turn is useful for identifying modulators of
CC   hu-Aspl alpha-secretase activity, where modulators that increase
CC   hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
CC   disease (AD) which causes progressive dementia with consequent
CC   formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC   neuronal loss. Hu-Aspl protease substrate is useful for assaying
CC   hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
CC   the substrate under acidic conditions and determining the level of
CC   hu-Aspl proteolytic activity. The present sequence is Aspl protein
CC   from human. Aspl gene is localised on chromosome 21.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     22              Gaps:           0

US-09-806-194-1 (1-1804) x AAE10628 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
```

```
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||  ||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTCGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 8
AAE10656
ID AAE10656 standard; Protein; 518 AA.
XX
AC AAE10656;
XX
DT 10-DEC-2001 (first entry)
XX
DE Human-Asp 1 protein lacking TM domain and containing (His)6 tag.
XX
KW Human; aspartyl protease 1; Aspl; amyloid precursor protein; APP;
KW Alzheimer's disease; AD; dementia; neurofibrillary tangle; gliosis;
KW amyloid plaque; neuronal loss; proteolytic; nootropic; neuroprotective.
XX
OS Homo sapiens.
OS Synthetic.
XX
PN GB2357767-A.
XX
PD 04-JUL-2001.
XX
PF 22-SEP-2000; 2000GB-0023315.
XX
PR 23-SEP-1999; 99US-0155493.
PR 23-SEP-1999; 99US-0404133.
PR 23-SEP-1999; 99WO-US20881.
PR 13-OCT-1999; 99US-0416901.
PR 06-DEC-1999; 99US-0169232.
XX
PA (PHAA) PHARMACIA & UPJOHN CO.
XX
PI Bienkowski MJ, Gurney M;
XX
DR WPI; 2001-444208/48.
XX
PT Polypeptide comprising fragments of human aspartyl protease with
PT amyloid precursor protein processing activity and alpha-secretase
PT activity, for identifying modulators useful in treating Alzheimer's
PT disease -
XX
PS Example 14; Page 155-156; 187pp; English.
XX
CC *The patent discloses human aspartyl protease 1 (hu-Aspl) or modified*
CC Aspl proteins which lack transmembrane domain or amino terminal
CC domain or cytoplasmic domain and retains alpha-secretase activity
CC and amyloid protein precursor (APP) processing activity. The proteins
CC of the invention are useful for assaying hu-Aspl alpha-secretase
CC activity, which in turn is useful for identifying modulators of
CC hu-Aspl alpha-secretase activity, where modulators that increase
CC hu-Aspl alpha-secretase activity are useful for treating Alzheimer's
CC disease (AD) which causes progressive dementia with consequent
CC formation of amyloid plaques, neurofibrillary tangles, gliosis and
CC *neuronal loss. Hu-Aspl protease substrate is useful for assaying*
CC hu-Aspl proteolytic activity, by contacting hu-Aspl protein with
CC the substrate under acidic conditions and determining the level of
CC hu-Aspl proteolytic activity. The present sequence is human Asp 1
CC protein lacking a transmembrane (TM) *domain and containing (His)*6
CC tag. This sequence is generated from human Asp 1 protein by the
CC deletion of its C-terminal TM domain and addition of hexa-histidine
CC tag at its C-terminus.
XX
SQ Sequence 518 AA;

Alignment Scores:

Pred. No.:	4.01e-213	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0

```
Best Local Similarity:  100.00%        Mismatches:    0
Query Match:            80.47%         Indels:        0
DB:                     22             Gaps:          0

US-09-806-194-1 (1-1804) x AAE10656 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTC 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGGCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGICTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 9
AAU29059
ID   AAU29059 standard; Protein; 518 AA.
XX
AC   AAU29059;
XX
DT   18-DEC-2001  (first entry)
XX
DE   Human PRO polypeptide sequence #36.
XX
KW   PRO polypeptide; mammal; tumour; cancer; human; cattle; horse; sheep;
KW   dog; cat; pig; goat; rabbit; tumour necrosis factor alpha; TNF-alpha;
KW   blood; chondrocyte cell; cell proliferation; cell differentiation; colon;
KW   adrenal; lung; breast; prostate; rectum; cervix; liver; genetic disorder.
XX
OS   Homo sapiens.
XX
PN   WO200168848-A2.
XX
PD   20-SEP-2001.
XX
PF   28-FEB-2001; 2001WO-US06520.
XX
PR   01-MAR-2000; 2000WO-US05601.
PR   02-MAR-2000; 2000WO-US05841.
PR   03-MAR-2000; 2000US-187202P.
PR   06-MAR-2000; 2000US-186968P.
PR   14-MAR-2000; 2000US-189320P.
PR   14-MAR-2000; 2000US-189328P.
PR   15-MAR-2000; 2000WO-US06884.
PR   21-MAR-2000; 2000US-190828P.
PR   21-MAR-2000; 2000US-191007P.
PR   21-MAR-2000; 2000US-191048P.
PR   21-MAR-2000; 2000US-191314P.
PR   28-MAR-2000; 2000US-192655P.
PR   29-MAR-2000; 2000US-193032P.
```

```
PR   29-MAR-2000; 2000US-193053P.
PR   30-MAR-2000; 2000WO-US08439.
PR   04-APR-2000; 2000US-194449P.
PR   04-APR-2000; 2000US-194647P.
PR   11-APR-2000; 2000US-195975P.
PR   11-APR-2000; 2000US-196000P.
PR   11-APR-2000; 2000US-196187P.
PR   11-APR-2000; 2000US-196690P.
PR   11-APR-2000; 2000US-196820P.
PR   18-APR-2000; 2000US-198121P.
PR   18-APR-2000; 2000US-198585P.
PR   25-APR-2000; 2000US-199397P.
PR   25-APR-2000; 2000US-199550P.
PR   25-APR-2000; 2000US-199654P.
PR   03-MAY-2000; 2000US-201516P.
PR   17-MAY-2000; 2000WO-US13705.
PR   22-MAY-2000; 2000WO-US14042.
PR   30-MAY-2000; 2000WO-US14941.
PR   02-JUN-2000; 2000WO-US15264.
PR   05-JUN-2000; 2000US-209832P.
PR   28-JUL-2000; 2000WO-US20710.
PR   22-AUG-2000; 2000US-0644848.
PR   24-AUG-2000; 2000WO-US23328.
PR   08-NOV-2000; 2000WO-US30952.
PR   01-DEC-2000; 2000WO-US32678.
PR   20-DEC-2000; 2000WO-US34956.
XX
PA   (GETH ) GENENTECH INC.
XX
PI   Baker KP, Chen J, Desnoyers L, Goddard A, Godowski PJ, Gurney AL;
PI   Pan J, Smith V, Watanabe CK, Wood WI, Zhang Z;
XX
DR   WPI; 2001-602746/68.
DR   N-PSDB; AAS45960.
XX
PT   Novel nucleic acids encoding PRO polypeptides, used to diagnose the
PT   presence of tumours, such as prostate and breast tumours, in mammals and
PT   to screen for modulators of the compounds -
XX
PS   Claim 11; Fig 72; 774pp; English.
XX
CC   Sequences AAU29024-AAU29328 represent PRO polypeptides of the invention.
CC   The PRO polypeptides and their associated nucleic acids can be used to
CC   detect the presence of a tumour in a mammal by comparing the level of
CC   expression of a PRO polypeptide in a test sample of cells from the animal
CC   and a control sample of normal cells, whereby a higher level of
CC   expression in the test sample indicates the presence of a tumour in the
CC   mammal. Mammals include dogs, cats, cattle, horses, sheep, pigs, goats
CC   and rabbits but are preferably human. The polypeptides can be used to
CC   stimulate tumour necrosis factor (TNF) alpha release from human blood,
CC   when contacted with it. A specific polypeptide can be used to stimulate
CC   the proliferation or differentiation of chondrocyte cells. The PRO
CC   proteins can be used to determine the presence of tumours and also
CC   susceptibility to tumour development, particularly adrenal, lung, colon,
CC   breast, prostate, rectal, cervical, or liver tumours, in mammalian
CC   subjects. The oligonucleotide probes specific for the PRO nucleic acids
CC   can be used for genetic analysis of individuals with genetic disorders.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     22              Gaps:           0

US-09-806-194-1 (1-1804) x AAU29059 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380
```

```
Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATATGAATGTTACCGATTCGGCATTCCCCA 1200
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1321
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 10
AAE06858
ID AAE06858 standard; Protein; 518 AA.
XX
AC AAE06858;
XX
DT 23-OCT-2001 (first entry)
XX
DE Human aspartyl protease 1 (Hu-Aspl) protein.
XX
KW Human; aspartyl protease 1; Asp 1; beta-amyloid precursor protein; APP;
KW beta-secretase; Alzheimer's disease; dementia; amyloid plaque; gliosis;
KW neurofibrillary tangle; neuronal loss; amyloid-beta peptide; nootropic;
KW neuroprotective; antisense therapy; gene therapy; chromosome 21.
XX
OS Homo sapiens.
XX
FH Key Location/Qualifiers
FT Peptide 1..20
FT /label= Signal_peptide
FT Protein 21..518
FT /note= "Mature human aspartyl protease 1 (Hu-Aspl)"
FT Domain 469..492
FT /label= Transmembrane_domain
XX
PN WO200150829-A2.
XX
PD 19-JUL-2001.
XX
PF 09-MAY-2001; 2001WO-IB00799.
XX
PR 09-MAY-2001; 2001WO-IB00799.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
DR WPI; 2001-483072/52.
DR N-PSDB; AAD13020.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity -
XX
PS Example 2; Fig 1; 185pp; English.
XX
CC The invention relates to human aspartyl proteases (Hu-Asp), beta-amyloid
CC precursor protein (APP) isoforms and their corresponding DNA molecules.
CC Human aspartyl proteases can act as beta-secretase proteases useful for
CC treating Alzheimer's disease. APP isoforms are useful for identifying
CC modulators of amyloid-beta peptide production, for use in designing
CC therapeutics for the treatment and prevention of Alzheimer's disease,
CC dementia, formation of amyloid plaques, neurofibrillary tangles, gliosis
CC and neuronal loss. APP isoforms are also used in methods for identifying
CC inhibitors and modulators of human Asp2 activity. The invention relates
CC to a method for identifying agents that modulate the activity of human
CC aspartyl protease Asp2. Amyloid-beta peptides obtained from APP are used
CC as a means to screen in cellular assays for the inhibitors of beta- and
CC gamma- secretase. Hu-Asp DNA fragments are useful as probes or primers in
CC polymerase chain reactions (PCR). The probes are useful for detecting
CC Hu-Asp nucleic acids in in vitro assays and in Northern and Southern
CC blots. The present sequence is human aspartyl protease 1 (Hu-Aspl).
CC Hu-Asp 1 gene is localised on chromosome 21.
XX
SQ Sequence 518 AA;

Alignment Scores:

Pred. No.:	4.01e-213	Length:	518
Score:	2687.00	Matches:	518
Percent Similarity:	100.00%	Conservative:	0
Best Local Similarity:	100.00%	Mismatches:	0
Query Match:	80.47%	Indels:	0
DB:	22	Gaps:	0

US-09-806-194-1 (1-1804) x AAE06858 (1-518)

```
Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
```

```
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 11
AAU06602
ID   AAU06602 standard; Protein; 518 AA.
XX
AC   AAU06602;
XX
DT   24-OCT-2001  (first entry)
XX
DE   Human Aspartyl protease 1 (Aspl).
XX
KW   Human; Aspartyl protease; Asp1; Asp2; beta-secretase; nootropic;
KW   neuroprotective; amyloid protein precursor; APP; Alzheimer's disease;
KW   amyloid-beta; Abeta.
XX
OS   Homo sapiens.
XX
PN   WO200149098-A2.
XX
PD   12-JUL-2001.
XX
PF   09-MAY-2001; 2001WO-IB00798.
XX
PR   09-MAY-2001; 2001WO-IB00798.
XX
PA   (BIEN/) BIENKOWSKI M J.
PA   (GURN/) GURNEY M E.
PA   (HEIN/) HEINRIKSON R L.
PA   (PARO/) PARODI L A.
PA   (YANR/) YAN R.
XX
PI   Bienkowski MJ,  Gurney ME,  Heinrikson RL,  Parodi LA,  Yan R;
XX
DR   WPI; 2001-502549/55.
DR   N-PSDB; AAS11516.
XX
PT   Novel purified polypeptide comprising fragment of mammalian aspartyl
PT   protease 2, lacking Asp2 transmembrane domain and retaining beta
PT   secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT   activity -
XX
PS   Example 2; Fig 1; 185pp; English.
XX
CC   The invention relates to a purified polypeptide comprising a fragment of
CC   mammalian aspartyl protease (Asp)2 protein which lacks the Asp2
CC   transmembrane domain and the Asp2 protein, and where the polypeptide and
CC   the fragment retain the beta-secretase activity of the mammalian Asp2
CC   protein. The invention also details polynucleotides for the Asp
CC   proteins and vectors expressing them, and a polypeptide (isoform of
CC   amyloid protein precursor (APP)) comprising the amino acid sequence of an
CC   APP or its fragment containing an APP cleavage site recognizable by a
CC   mammalian beta-secretase, and further comprising two lysine residues at
CC   the carboxyl terminus of the amino acid sequence of the mammalian APP or
CC   APP fragment. Also included in the invention are methods of identifying
CC   modulators or inhibitors of Asp2. Modulators and inhibitors of Asp2 are
CC   useful for treating Alzheimer's disease. APP is useful in methods for
CC   identifying inhibitors or modulators of human Asp2 activity and
CC   amyloid-beta (Abeta) peptide production. APP is also useful in designing
CC   therapeutics for the treatment or prevention of Alzheimer's disease.
CC   APP comprising the APP-Sw-beta-secretase peptide sequence (NLDA), which
CC   is associated with increased levels of Abeta processing is useful in
CC   assays relating the Alzheimer's research. The expression vector is useful
CC   for recombinantly expressing APP. Nucleic acids that hybridise to
CC   Asp oligonucleotides are  useful as probes or primers. The probes are
CC   useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC   Northern and Southern blots. The present sequence is human Asp1.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213       Length:         518
Score:                  2687.00         Matches:        518
Percent Similarity:     100.00%         Conservative:   0
Best Local Similarity:  100.00%         Mismatches:     0
Query Match:            80.47%          Indels:         0
DB:                     22              Gaps:           0

US-09-806-194-1 (1-1804) x AAU06602 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
```

```
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTC 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyIle 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTCTGTATAAA 780
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140

Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518
```

RESULT 12
AAU07201
ID AAU07201 standard; Protein; 518 AA.
XX
AC AAU07201;
XX
DT 24-OCT-2001 (first entry)
XX
DE Human aspartyl protease 1 (Asp-1).
XX
KW Human; aspartyl protease 1; Asp-1; nootropic; neuroprotective;
KW aspartyl protease 2; Asp2; amyloid protein precursor; APP;
KW beta-secretase; Alzheimer's disease.
XX
OS Homo sapiens.
XX
PN WO200149097-A2.
XX
PD 12-JUL-2001.
XX
PF 09-MAY-2001; 2001WO-IB00797.
XX
PR 09-MAY-2001; 2001WO-IB00797.
XX
PA (BIEN/) BIENKOWSKI M J.
PA (GURN/) GURNEY M E.
PA (HEIN/) HEINRIKSON R L.
PA (PARO/) PARODI L A.
PA (YANR/) YAN R.
XX
PI Bienkowski MJ, Gurney ME, Heinrikson RL, Parodi LA, Yan R;
XX
DR WPI; 2001-502548/55.
DR N-PSDB; AAS11701.
XX
PT Novel purified polypeptide comprising fragment of mammalian aspartyl
PT protease 2, lacking Asp2 transmembrane domain and retaining beta
PT secretase activity of Asp2 useful for identifying inhibitors of Asp2
PT activity -
XX
PS Example 2; Fig 1; 185pp; English.
XX
CC The invention relates to a novel purified polypeptide comprising a

```
CC   fragment of mammalian aspartyl protease 2 (Asp2) protein which lacks the
CC   Asp2 transmembrane domain and the Asp2 protein, and where the polypeptide
CC   and the fragment retain the beta-secretase activity of the mammalian Asp2
CC   protein. Also included is an isoform of amyloid protein precursor (APP)
CC   comprising the amino acid sequence of a APP or its fragment containing
CC   an APP cleavage site recognisable by a mammalian beta-secretase, and
CC   further comprising two lysine residues at the carboxyl terminus of the
CC   amino acid sequence of the mammalian APP or APP fragment. The
CC   polypeptides are used for assaying for modulators of beta-secretase
CC   activity; identifying agents that inhibit the APP processing activity
CC   of human Asp2 aspartyl protease (Hu-Asp2); identifying agents that
CC   modulate the activity of Asp2; and for reducing cellular production of
CC   amyloid beta (Abeta) from APP. Agents identified by the above methods
CC   are useful for treating Alzheimer's disease; and for identifying
CC   modulators of amyloid-beta (Abeta) peptide production, for use in
CC   designing therapeutics for the treatment or prevention of Alzheimer's
CC   disease. Probes and primers derived from Asp nucleic acid sequences
CC   are useful for detecting Hu-Asp nucleic acids in in vitro assays and in
CC   Northern and Southern blots. The present sequence represents the
CC   amino acid sequence of human Asp-1.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:                4.01e-213      Length:          518
Score:                    2687.00        Matches:         518
Percent Similarity:       100.00%        Conservative:    0
Best Local Similarity:    100.00%        Mismatches:      0
Query Match:              80.47%         Indels:          0
DB:                       22             Gaps:            0

US-09-806-194-1 (1-1804) x AAU07201 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGGCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGGCAGGA 360
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCITGGGTGGAATTGAACCAAGTTTGTATAAA 780
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 13
AAE02580
ID   AAE02580 standard; Protein; 518 AA.
XX
AC   AAE02580;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human aspartyl protease 1 (Asp 1).
```

```
XX
KW   Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW   Alzheimer's disease; antialzheimer's; aspartyl protease 1; Asp 1;
KW   beta-secretase; chromosome 21.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Peptide         1..20
FT                   /label= Signal_peptide
FT   Peptide         22..62
FT                   /label= Asp_1_prepropeptide
FT   Peptide         23..62
FT                   /label= Asp_1_propeptide
FT   Protein         63..518
FT                   /label= Mature_human_Asp_1_protein
FT                   /note= "Specifically claimed"
FT   Active-site     87..89
FT                   /label= Active_site_1
FT   Active-site     110..113
FT                   /label= Active_site_2
FT   Active-site     303..305
FT                   /label= Active_site_3
FT   Domain          469..492
FT                   /label= Transmembrane_domain
FT   Domain          493..518
FT                   /label= Cytoplasmic_domain
FT   Region          497..518
FT                   /note= "Peptide #1"
XX
PN   WO200123533-A2.
XX
PD   05-APR-2001.
XX
PF   22-SEP-2000; 2000WO-US26080.
XX
PR   23-SEP-1999;   99US-0155493.
PR   23-SEP-1999;   99WO-US20881.
PR   13-OCT-1999;   99US-0416901.
PR   06-DEC-1999;   99US-0169232.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney M,  Bienkowski MJ;
XX
DR   WPI; 2001-290516/30.
DR   N-PSDB; AAD06738.
XX
PT   Enzymes that cleave the alpha-secretase site of the amyloid precursor
PT   protein, useful for the treatment of Alzheimer's disease -
XX
PS   Claim 29; Fig 1; 189pp; English.
XX
CC   The present invention relates to enzymes for cleaving the alpha-
CC   secretase site of the amyloid precursor protein (APP) and methods of
CC   identifying those enzymes. The methods may be used to identify enzymes
CC   that may be used to cleave the alpha-secretase cleavage site of the APP
CC   protein. The enzymes may be used to treat or modulate the progress of
CC   Alzheimer's disease. The present sequence is human aspartyl protease 1
CC   (Asp 1). Asp 1 has alpha-secretase protease and beta-secretase
CC   protease activities. Asp 1 gene is located on chromosome 21.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:              4.01e-213      Length:          518
Score:                  2687.00        Matches:         518
Percent Similarity:     100.00%        Conservative:    0
Best Local Similarity:  100.00%        Mismatches:      0
Query Match:            80.47%         Indels:          0
DB:                     22             Gaps:            0

US-09-806-194-1 (1-1804) x AAE02580 (1-518)

Qy        1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db        1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy       61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db       21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy      121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db       41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy      181 GCGCTCGCCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db       61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy      241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db       81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy      301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db      101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy      361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db      121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy      421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db      141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy      481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db      161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy      541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db      181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy      601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db      201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy      661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db      221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy      721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db      241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260

Qy      781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
Db      261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy      841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
Db      281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy      901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGGTG 960
Db      301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy      961 GAAGCTGTGGCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
Db      321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy     1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
Db      341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360
```

```
Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCAGAATCCTGCCTCAGCTTTAC 1140
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCGGCT 1380
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 14
AAE02608
ID   AAE02608 standard; Protein; 518 AA.
XX
AC   AAE02608;
XX
DT   10-AUG-2001  (first entry)
XX
DE   Human Aspartyl protease-1 (Asp-1) deltaTM (His)6 protein.
XX
KW   Human; alpha-secretase; amyloid precursor protein; APP; therapy;
KW   Alzheimer's disease; antialzheimer's; aspartyl protease 1; Asp1;
KW   beta-secretase; Asp-1 deltaTM (His)6 protein.
XX
OS   Homo sapiens.
OS   Synthetic.
XX
PN   WO200123533-A2.
XX
PD   05-APR-2001.
XX
PF   22-SEP-2000; 2000WO-US26080.
XX
PR   23-SEP-1999;   99US-0155493.
PR   23-SEP-1999;   99WO-US20881.
PR   13-OCT-1999;   99US-0416901.
PR   06-DEC-1999;   99US-0169232.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Gurney M,  Bienkowski MJ;
XX
DR   WPI; 2001-290516/30.
XX
PT   Enzymes that cleave the alpha-secretase site of the amyloid precursor
PT   protein, useful for the treatment of Alzheimer's disease -
XX
PS   Example 14; Page 183-184; 189pp; English.
XX
CC   The present invention relates to enzymes for cleaving the alpha-
CC   secretase site of the amyloid precursor protein (APP) and methods of
CC   identifying those enzymes. The methods may be used to identify enzymes
CC   that may be used to cleave the alpha-secretase cleavage site of the APP
CC   protein. The enzymes may be used to treat or modulate the progress of
CC   Alzheimer's disease. The present sequence is human Aspartyl protease-1
CC   (Asp-1) deltaTM (His)6 protein which is used for the expression of
CC   pre-pro-human-Aspartyl protease 1 (Aspl). This protein is obtained by
CC   replacing C-terminal transmembrane and cytoplasmic domains with a
CC   hexahistidine purification tag in the human Aspartyl protease 1.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:                4.01e-213       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       22              Gaps:            0

US-09-806-194-1 (1-1804) x AAE02608 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCTCCGGGTGGCCGCGGCCACGAAC 120
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
Db         81 ValAspAsnLeuGlnGlyAspSerGlyArgGlyTyrTyrLeuGluMetLeuIleGlyThr 100

Qy        301 CCCCCGCAGAAGCTACAGATTCTCGTTGACACTGGAAGCAGTAACTTTGCCGTGGCAGGA 360
Db        101 ProProGlnLysLeuGlnIleLeuValAspThrGlySerSerAsnPheAlaValAlaGly 120

Qy        361 ACCCCGCACTCCTACATAGACACGTACTTTGACACAGAGAGGTCTAGCACATACCGCTCC 420
Db        121 ThrProHisSerTyrIleAspThrTyrPheAspThrGluArgSerSerThrTyrArgSer 140

Qy        421 AAGGGCTTTGACGTCACAGTGAAGTACACACAAGGAAGCTGGACGGGCTTCGTTGGGGAA 480
Db        141 LysGlyPheAspValThrValLysTyrThrGlnGlySerTrpThrGlyPheValGlyGlu 160

Qy        481 GACCTCGTCACCATCCCCAAAGGCTTCAATACTTCTTTTCTTGTCAACATTGCCACTATT 540
Db        161 AspLeuValThrIleProLysGlyPheAsnThrSerPheLeuValAsnIleAlaThrIle 180

Qy        541 TTTGAATCAGAGAATTTCTTTTTGCCTGGGATTAAATGGAATGGAATACTTGGCCTAGCT 600
Db        181 PheGluSerGluAsnPhePheLeuProGlyIleLysTrpAsnGlyIleLeuGlyLeuAla 200

Qy        601 TATGCCACACTTGCCAAGCCATCAAGTTCTCTGGAGACCTTCTTCGACTCCCTGGTGACA 660
Db        201 TyrAlaThrLeuAlaLysProSerSerSerLeuGluThrPhePheAspSerLeuValThr 220

Qy        661 CAAGCAAACATCCCCAACGTTTTCTCCATGCAGATGTGTGGAGCCGGCTTGCCCGTTGCT 720
Db        221 GlnAlaAsnIleProAsnValPheSerMetGlnMetCysGlyAlaGlyLeuProValAla 240

Qy        721 GGATCTGGGACCAACGGAGGTAGTCTTGTCTTGGGTGGAATTGAACCAAGTTTGTATAAA 780
Db        241 GlySerGlyThrAsnGlyGlySerLeuValLeuGlyGlyIleGluProSerLeuTyrLys 260
```

```
Qy        781 GGAGACATCTGGTATACCCCTATTAAGGAAGAGTGGTACTACCAGATAGAAATTCTGAAA 840
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        261 GlyAspIleTrpTyrThrProIleLysGluGluTrpTyrTyrGlnIleGluIleLeuLys 280

Qy        841 TTGGAAATTGGAGGCCAAAGCCTTAATCTGGACTGCAGAGAGTATAACGCAGACAAGGCC 900
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        281 LeuGluIleGlyGlyGlnSerLeuAsnLeuAspCysArgGluTyrAsnAlaAspLysAla 300

Qy        901 ATCGTGGACAGTGGCACCACGCTGCTGCGCCTGCCCCAGAAGGTGTTTGATGCGGTGCTG 960
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        301 IleValAspSerGlyThrThrLeuLeuArgLeuProGlnLysValPheAspAlaValVal 320

Qy        961 GAAGCTGTGGCCCCGCGCATCTCTGATTCCAGAATTCTCTGATGGTTTCTGGACTGGGTCC 1020
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        321 GluAlaValAlaArgAlaSerLeuIleProGluPheSerAspGlyPheTrpThrGlySer 340

Qy       1021 CAGCTGGCGTGCTGGACGAATTCGGAAACACCTTGGTCTTACTTCCCTAAAATCTCCATC 1080
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        341 GlnLeuAlaCysTrpThrAsnSerGluThrProTrpSerTyrPheProLysIleSerIle 360

Qy       1081 TACCTGAGAGATGAGAACTCCAGCAGGTCATTCCGTATCACAATCCTGCCTCAGCTTTAC 1140
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        361 TyrLeuArgAspGluAsnSerSerArgSerPheArgIleThrIleLeuProGlnLeuTyr 380

Qy       1141 ATTCAGCCCATGATGGGGGCCGGCCTGAATTATGAATGTTACCGATTCGGCATTTCCCCA 1200
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        381 IleGlnProMetMetGlyAlaGlyLeuAsnTyrGluCysTyrArgPheGlyIleSerPro 400

Qy       1201 TCCACAAATGCGCTGGTGATCGGTGCCACGGTGATGGAGGGCTTCTACGTCATCTTCGAC 1260
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        401 SerThrAsnAlaLeuValIleGlyAlaThrValMetGluGlyPheTyrValIlePheAsp 420

Qy       1261 AGAGCCCAGAAGAGGGTGGGCTTCGCAGCGAGCCCCTGTGCAGAAATTGCAGGTGCTGCA 1320
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        421 ArgAlaGlnLysArgValGlyPheAlaAlaSerProCysAlaGluIleAlaGlyAlaAla 440

Qy       1321 GTGTCTGAAATTTCCGGGCCTTTCTCAACAGAGGATGTAGCCAGCAACTGTGTCCCCGCT 1380
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        441 ValSerGluIleSerGlyProPheSerThrGluAspValAlaSerAsnCysValProAla 460

Qy       1381 CAGTCTTTGAGCGAGCCCATTTTGTGGATTGTGTCCTATGCGCTCATGAGCGTCTGTGGA 1440
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        461 GlnSerLeuSerGluProIleLeuTrpIleValSerTyrAlaLeuMetSerValCysGly 480

Qy       1441 GCCATCCTCCTTGTCTTAATCGTCCTGCTGCTGCTGCCGTTCCGGTGTCAGCGTCGCCCC 1500
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        481 AlaIleLeuLeuValLeuIleValLeuLeuLeuLeuProPheArgCysGlnArgArgPro 500

Qy       1501 CGTGACCCTGAGGTCGTCAATGATGAGTCCTCTCTGGTCAGACATCGCTGGAAA 1554
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db        501 ArgAspProGluValValAsnAspGluSerSerLeuValArgHisArgTrpLys 518

RESULT 15
ABB78589
ID   ABB78589 standard; Protein; 518 AA.
XX
AC   ABB78589;
XX
DT   16-JUL-2002  (first entry)
XX
DE   Human Asp-1 protein sequence SEQ ID NO:2.
XX
KW   Human; Asp-1; Asp-2; aspartyl protease; enzyme; Alzheimer's disease;
KW   proteolytic; chromosome 21.
XX
OS   Homo sapiens.
XX
PN   GB2367060-A.
XX
PD   27-MAR-2002.
XX
PF   29-OCT-2001; 2001GB-0025934.
XX
PR   23-SEP-1999;   99US-155493P.
PR   23-SEP-1999;   99US-0404133.
PR   23-SEP-1999;   99WO-US20881.
PR   13-OCT-1999;   99US-0416901.
PR   06-DEC-1999;   99US-169232P.
PR   22-SEP-2000; 2000GB-0023315.
XX
PA   (PHAA ) PHARMACIA & UPJOHN CO.
XX
PI   Bienkowkski MJ,  Gurney M;
XX
DR   WP1; 2002-396337/43.
DR   N-PSDB; ABL52456.
XX
PT   Human aspartyl protease 1 substrates useful in assays to detect
PT   aspartyl protease activity, e.g. for the diagnosis of Alzheimer's
PT   disease -
XX
PS   Claim 7; Fig 1; 182pp; English.
XX
CC   The present invention describes a human aspartyl protease 1 (hu-Asp1)
CC   substrate (I) which comprises a peptide of no more than 50 amino acids,
CC   and which comprises the 8 amino acid sequence Gly-Leu-Ala-Leu-Ala-Leu-
CC   Glu-Pro. Also described are: (1) a method (II) for assaying hu-Asp1
CC   proteolytic activity, comprising: (a) contacting a hu-Asp1 protein with
CC   (I) under acidic conditions; and (b) determining the level of hu-Asp1
CC   proteolytic activity; (2) a purified polynucleotide (III) comprising a
CC   nucleotide sequence that hybridises under stringent conditions to the
CC   non-coding strand complementary to a defined 1804 nucleotide sequence
CC   (see ABL52456) where the nucleotide sequence encodes a polypeptide having
CC   Asp1 proteolytic activity and lacks nucleotides encoding a transmembrane
CC   domain); (3) a purified polynucleotide (III') comprising a sequence that
CC   hybridises under stringent conditions to (III) (the nucleotide sequence
CC   encodes a polypeptide further lacking a pro-peptide domain corresponding
CC   to amino acids 23-62 of hu-Asp1 (see ABB78589)); (4) a vector (IV)
CC   comprising (III) or (III'); and (5) a host cell (V) transformed or
CC   transfected with (III), (III') and/or (IV). The hu-Asp1 protease
CC   substrate (I) may be used as an enzyme substrate in assays to detect
CC   aspartyl protease activity, (II) and therefore diagnose diseases
CC   associated with aberrant hu-Asp1 expression and activity such as
CC   Alzheimer's disease. Hu-Asp1 has been localised to chromosome 21, while
CC   hu-Asp2 has been localised to chromosome 11q23.3-24.1. The present
CC   sequence represents hu-Asp1 from the present invention.
XX
SQ   Sequence   518 AA;

Alignment Scores:
Pred. No.:                4.01e-213       Length:          518
Score:                    2687.00         Matches:         518
Percent Similarity:       100.00%         Conservative:    0
Best Local Similarity:    100.00%         Mismatches:      0
Query Match:              80.47%          Indels:          0
DB:                       23              Gaps:            0

US-09-806-194-1 (1-1804) x ABB78589 (1-518)

Qy          1 ATGGGCGCACTGGCCCGGGCGCTGCTGCTGCCTCTGCTGGCCCAGTGGCTCCTGCGCGCC 60
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db          1 MetGlyAlaLeuAlaArgAlaLeuLeuLeuProLeuLeuAlaGlnTrpLeuLeuArgAla 20

Qy         61 GCCCCGGAGCTGGCCCCCGCGCCCTTCACGCTGCCCCCTCCGGGTGGCCGCGGCCACGAAC 120
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         21 AlaProGluLeuAlaProAlaProPheThrLeuProLeuArgValAlaAlaAlaThrAsn 40

Qy        121 CGCGTAGTTGCGCCCACCCCGGGACCCGGGACCCCTGCCGAGCGCCACGCCGACGGCTTG 180
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         41 ArgValValAlaProThrProGlyProGlyThrProAlaGluArgHisAlaAspGlyLeu 60

Qy        181 GCGCTCGCCCTGGAGCCTGCCCTGGCGTCCCCCGCGGGCGCCGCCAACTTCTTGGCCATG 240
              ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db         61 AlaLeuAlaLeuGluProAlaLeuAlaSerProAlaGlyAlaAlaAsnPheLeuAlaMet 80

Qy        241 GTAGACAACCTGCAGGGGGACTCTGGCCGCGGCTACTACCTGGAGATGCTGATCGGGACC 300
```

Search completed: September 25, 2003, 11:00:02
Job time : 146.5 secs